UNIVERSIDADE FEDERAL DOS VALES DO JEQUITINHONHA E MUCURI
Programa de Pós-Graduação em Reabilitação e Desempenho Funcional

Keysy Karoline Souza Andrade

Qualidade das propriedades de medidas dos instrumentos de avaliação funcional de pessoas com distrofias musculares: uma revisão sistemática

Diamantina
2021

Keysy Karoline Souza Andrade

Qualidade das propriedades de medidas dos instrumentos de avaliação funcional de pessoas com distrofias musculares: uma revisão sistemática

Dissertação apresentada ao programa de Pós-Graduação em Reabilitação e Desempenho Funcional da Universidade Federal dos Vales do Jequitinhonha e Mucuri, como requisito para obtenção do título de Mestre.

Orientadora: Profa. Dra. Thaís Peixoto Gaiad
Coorientador: Prof. Dr. Hércules Ribeiro Leite

Diamantina

2021

Ficha catalográfica

Catalogação na fonte - Sisbi/UFVJM

A553 Andrade, Keysy Karoline Souza
2021 Qualidade das propriedades de medida dos instrumentos de avaliação funcional de pessoas com distrofias musculares: uma revisão sistemática [manuscrito] / Keysy Karoline Souza Andrade. -- Diamantina, 2021.
59 p. : il.

Orientador: Prof. Thaís Peixoto Gaiad.
Coorientador: Prof. Hércules Ribeiro Leite.

Dissertação (Mestrado em Reabilitação e Desempenho Funcional) -- Universidade Federal dos Vales do Jequitinhonha e Mucuri, Programa de Pós-Graduação em Reabilitação e Desempenho Funcional, Diamantina, 2021.

1. Distrofias musculares. 2. CIF. 3. Doenças neuromusculares. 4. Instrumentos de avaliação. I. Gaiad, Thaís Peixoto . II. Leite, Hércules Ribeiro . III. Universidade Federal dos Vales do Jequitinhonha e Mucuri. IV. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFVJM
com os dados fornecidos pelo(a) autor(a).
Bibliotecário Rodrigo Martins Cruz / CRB6-2886
Técnico em T.I. Thales Francisco Mota Carvalho

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DOS VALES DO JEQUITINHONHA E MUCURI

KEYSY KAROLINE SOUZA ANDRADE

Qualidade das propriedades de medidas dos instrumentos de avaliação funcional de pessoas com distrofias musculares: uma revisão sistemática

Dissertação apresentada ao programa de Pós-Graduação em REABILITAÇÃO E DESEMPENHO FUNCIONAL da Universidade Federal dos Vales do Jequitinhonha e Mucuri, **nível de Mestrado,** como requisito parcial para obtenção do título de **Mestra** EM REABILITAÇÃO E DESEMPENHO FUNCIONAL.

Orientador: Profª **Thais Peixoto Gaiad Machado**

Data de aprovação 12/03/2021

Profa. Dr.ª THAIS PEIXOTO GAIAD MACHADO - UFVJM

Profa. Dr.ª VANESSA PEREIRA DE LIMA - UFVJM

Profa. Dr.ª AYRLES SILVA GONÇALVES BARBOSA MENDONÇA - UFAM

Documento assinado eletronicamente por **Thais Peixoto Gaiad Machado**, **Servidor**, em 17/03/2021, às 11:30, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Vanessa Pereira de Lima**, **Servidor**, em 17/03/2021, às 14:23, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Ayrles Silva Gonçalves Barbosa Mendonça**, **Usuário Externo**, em 17/03/2021, às 18:55, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Dedicatória

Dedico ao meu tio Renato Henry que me deu forças desde o inicio da minha trajetória acadêmica e que nunca deixou de acreditar em mim, mas pela vontade de Deus, se foi desse mundo para um mundo Celestial.

Agradecimentos

Agradeço primeiramente a Deus, por ter me dado forças.

Agradeço à minha mãe Sheila Sue que muitas vezes segurou as pontas no Studio Terapia do Toque e compreendeu as minhas ausências.

À minha Avó Marli Souza por participar da minha vida acadêmica, incentivando-me sempre em todas as minhas escolhas, conselhos valorosos, compreensão, amor eterno e por ser meu porto seguro.

À minha amiga Patrícia Campos pelo apoio, orações feitas e por ter trilhado mais um caminho ao meu lado desde a graduação juntas.

À best friend Alominha Spin que sempre esteve presente ao meu lado me apoiando na realização deste trabalho.

Agradecimentos especiais devem ser dados à professora Dra. Thais Peixoto Gaiad, pela confiança depositada, pelo acompanhamento pontual, paciência e competência demonstrada em cada detalhe. Na nossa segunda reunião colocou um leque de opções para que meu projeto fosse realizado da melhor maneira, entre laboratórios e online com a Revisão Sistemática. Então escolhi essa última com o tema que mais me chamava atenção: Distrofias Musculares e desde então seguimos juntas, fazendo reuniões nas semanas das disciplinas condensadas na UFVJM. Thaís, sou imensamente grata por ter feito muito mais do que imaginava.

Ao professor Hercules que me presenteou com a excelente professora Thais Gaiad, ele que nas reuniões, tanto pessoalmente quanto online tem uma simplicidade motivadora de cientista veloz.

Sou grata à Luana Soares que desde o início da extração de dados estava disposta a ficar horas e horas trabalhando de forma online, chamada de vídeo, ligação por muitos finais de semana até altas horas da noite, ela que não mediu esforços para se adaptar a minha rotina de trabalho que durante a semana só tinha tempo 21h da noite para bater os dados e no final vencemos. Sou grata, ainda, ao Caik e Nel que também estavam dispostos para que pudessemos resolver algo sobre a pesquisa.

Grata ao Ricardo que sua ajuda foi essencial para a realização desta pesquisa e pelas possibilidades de aprendizado criadas ao longo desta pesquisa.

E a todos aqueles que de alguma forma contribuíram e torceram pela concretização desta conquista.

RESUMO

As Doenças neuromusculares são condições que afetam os músculos e o sistema nervoso periférico, apresentando comprometimento progressivo em relação à função motora e pulmonar. As distrofias musculares - DM são um grupo de doenças neuromusculares de característica degenerativas hereditárias associadas à fraqueza muscular progressiva e degeneração das fibras musculares esqueléticas. Os instrumentos de avaliação são importantes para que sejam analisadas de forma criteriosa todas as alterações decorrentes da doença do indivíduo. A Classificação Internacional de Funcionalidade, Incapacidade e Saúde - CIF é uma ferramenta útil que descreve a funcionalidade e a incapacidade relacionadas às condições de saúde. Desse modo, o presente projeto trata-se de uma revisão sistemática que irá identificar os instrumentos validados para as distrofias musculares categorizando de acordo com os elementos da CIF. Para tanto serão utilizados artigos de validação de instrumentos para DM, nos bancos de dados eletrônicos: PUBmed, Lilacs, PEDro, Medline, Cochrane Lybrary, Scielo e Periodic Capes, no período de julho a novembro de 2019. Como resultados: Espera-se a obtenção de um artigo que facilite o acesso para profissionais da saúde, em especial aos fisioterapeutas, que aborde os instrumentos validados para distrofias musculares que contemplem os core sets da CIF, servindo de base para futuras intervenções.

Palavras-chave: CIF, Distrofias musculares, Doenças neuromusculares, Instrumentos de avaliação.

LISTA DE ILUSTRAÇÕES

LISTA DE TABELAS

LISTA DE SIGLAS

6MWT - 6 minute walking test

ABILHAND questionnaire

ACTIVILIM questionnaire

CAPE/PAC - Children's Assessment of Participation and Enjoyment and Preferences for Activities of Children

CFI - Comparative Fit Index

CIQ - Community Integration Questionnaire

COSMIN - Consensus-based Standards for the Selection of Health Status Measurement Instruments

DM1-Activ - Myotonic dystrophy type 1 activity and participation scale

DM1-Activ^{C} – DM1 activity and participation scale for clinical use

DMD SAT - Functional Ability Self-Assessment Tool

DMD Upper Limb PROM - Patient-reported outcome measure for upper limb function in Duchenne muscular dystrophy

EK - Egen Klassifikation Scale

EK2 - Egen Klassifikation Scale 2

FES-DMD-D1- Functional Evaluation Scale for Duchenne muscular dystrophy Go up and sit down on a chair

FES-DMD-D2 - Get up and sit on the floor

FES-DMD-D3 - Go up and down stairs

FES-DMD-D4 - Assessment of the walking activity

FSHD-COM - FSHD composite outcome measure

GMFM for FSHD - Motor function measure for Fukuyama congenital muscular dystrophy

GRADE - Grading of Recommendations Assessment, Development, and Evaluation

HGS - Hand grip strength

ICC - Intraclass correlation coefficient

LIFE-H - Assessment of Life Habits

MD - Muscular dystrophies

MDC - minimal detectible change

MDFRS - Dystrophy-specific functional rating scale

MDHI - Myotonic Dystrophy Health Index

MDSQ - Muscular Dystrophy Spine Questionnaire

MFM - Motor Function Measure

MFM-20 - Short-version MFM

MIRS - Muscular impairment rating scale

MRC - Medical Research Council

NDS - Neurological disability score

NMS - Neuromuscular-Score

NSAA - North Star Ambulatory Assessment

OMI - Outcome measurement instruments

PEM-CY - Participation and Environment Measure for Children and Youth

PUL - Performance of the Upper Limb

PROM – Patient-reported outcome measure

RMSEA - Root Mean Square Error of Approximation

SDC - Smallest detectable change

SEM - Standard Error of Measurement

SRMR - Standardized Root Mean Square Residuals

TLI - Tucker-Lewis Index

LISTA DE SÍMBOLOS

α - Cronbach’s alpha

k - Kappa index

SUMÁRIO

1. REVISÃO DE LITERATURA

1.1 Distrofias Musculares

Doenças neuromusculares são distúrbios caracterizados por alterações que envolvem os neurônios do corno anterior da medula, a raiz nervosa, nervos periféricos, a junção neuromuscular e/ou o músculo. Apresentam diferentes etiologias, sendo comum o comprometimento respiratório, cardiovascular e autonômico (Diniz, Lasmar, & Giannetti, 2010; Pereira, Castro, & Brochado, 2016).

As distrofias musculares são um grupo de doenças neuromusculares de característica degenerativas hereditárias associadas à fraqueza muscular progressiva e degeneração de fibras musculares esqueléticas. Podem ser transmitidas como traços autossômicos dominantes, autossômicos recessivos ou ligados ao X. A apresentação precoce durante a infância é geralmente associada a um fenótipo mais grave. O diagnóstico é baseado em características clínicas e patológicas; a maioria das distrofias musculares é classificada com base na confirmação genética molecular (Mah et al., 2015; Pértile, Almeida, Schlindwein-Zanini, Fernandes, & Helegda, 2014).

Manifestações clínicas das doenças neuromusculares são fraqueza muscular progressiva, atrofia muscular, problemas de deglutição e respiração. Esses sintomas podem levar a vulnerabilidade, com grande impacto no estado geral de saúde e vida cotidiana, com prováveis limitações em termos de tarefas ou participação na vida social em relação à habitação, trabalho e renda (Bos et al., 2013).

1.2. Instrumentos de avaliação funcional

Os instrumentos de avaliação são importantes para que sejam analisadas de forma criteriosa todas as alterações decorrentes da doença do individuo. Quando esse instrumento é padronizado e validado cientificamente torna-se benéfico para que possa ser reproduzido de forma fidedigna. As escalas mais utilizadas em pesquisas brasileiras para distrofias musculares são: Escala de avaliação Motora Funcional de Egen Klassification (EK), motor Function Measure (MFM); Escala Hammersmith Functional Motor, Escala de avaliação ambulatorial North Star (NSAA), Brooke Scale (Barra & Baraldi, 2013).

Instrumentos são considerados validados quando garantem a qualidade de seus resultados em diferentes momentos (Barra & Baraldi, 2013). Para garantir a qualidade nos

resultados de um instrumento, é necessário conhecer minuciosamente: itens, domínios, formas de avaliação e, principalmente, as propriedades de medida. A qualidade da informação dos instrumentos depende em parte de suas propriedades de medidas(Barra & Baraldi, 2013).

1.3. Propriedades de medida

Psicometria é o método que visa à mensuração e avaliação psicológica dos construtos subjetivos por meio de escalas, testes e questionários padronizados, denominados “medida psicométrica” (Cunha, Almeida Neto, & Stackfleth, 2017). As propriedades de medidas mais avaliadas são validade e a confiabilidade do instrumento. A validade é a qualidade de um instrumento para medir o construto para o qual foi construído, já a confiabilidade é o grau em que um instrumento permite a reprodução e consistência de resultados (Souza et al., 2017). Além dessas propriedades, tem-se o modelo trinário composto pela validade de construto, validade de conteúdo e validade de critério, que são satisfatórios na constatação da validade de um instrumento(Cunha, Almeida Neto, & Stackfleth, 2017).

O Consensus-based Satandards for the selection of Health Measureament Instruments (COSMIN) avalia as propriedades de medida de saúde, a fim de identificar os instrumentos de medida de saúde (Guanilo, 2017). Outcome Measuremen Instrument (OMI) é o instrumento de medida do resultados que refere-se como resultado medido, ou seja, é uma ferramenta para medir a qualidade ou quantidade dos resultados, a ferramenta pode ser um questionário, exames físicos, laboratoriais e imagens. (Prinsen, 2016).

O COSMIN contém uma lista de avaliação de 10 itens: validade de conteúdo, validade estrutural, consistência interna, validade transcultural, confiabilidade, erro de medição, validade de critério, teste de hipótese e capacidade de resposta.

Sendo categorizado em três domínios amplos: **Validade** incluindo validade de conteúdo/face, validade de critério, validade estrutural, teste de hipótese e validade transcultural. **Confiabilidade** contém três aspectos importantes a consistência interna, confiabilidade e erro de medição. A **Responsividade** que contém a propriedade do instrumento medido. A interpretabilidade também é um aspecto importante de um intrumento de medida, porém, não é considerada uma propriedade de medida (Echevarría-Guanilo, 2017; Terwee, 2017) (Figura 1).

Validade de conteúdo é o grau em que o conteúdo de um PROM são um reflexo adequado da construção a ser medido, ou seja, é a avaliação do tamanho da amostra de itens que representam o conteúdo ou campo definido. Validade de critério é o grau em que as pontuações de um PROM são um reflexo adequado de um 'padrão ouro’, ou seja, quando seus

escores são iguais aos escores do critério escolhido o construto é considerado válido. Validade transcultural é o grau em que o desempenho dos itens em um PROM traduzido ou adaptado culturalmente é um reflexo adequado do desempenho dos itens da versão original do PROM. Validade estrutural é grau em que as pontuações de um PROM é um reflexo adequado da dimensionalidade do construto a ser medido. Validade de constructo é o grau em que as pontuações de um PROM são consistentes com hipóteses (por exemplo, no que diz respeito a relações internas, relações com pontuações de outros instrumentos ou diferenças entre grupos relevantes) com base na suposição de que o PROM mede validamente a construção para ser medido (Prinsen, 2018; Souza, 2017).

Consistência interna é o grau de inter-relação entre os itens, ou seja, o grau em que os projetos medem os mesmos atributos e produzem resultados consistentes. Confiabilidade é o grau em que a medida é livre de erro de medição. Erro de medição é o erro sistemático e aleatório da pontuação de um paciente que não é atribuído a mudanças verdadeiras no construto a ser medido, se não houver erros na medição ou minimizados, a medição será considerada confiável. Responsividade é a capacidade de um PROM detectar mudanças ao longo do tempo no construto a ser medido (Prinsen, 2018; Echevarría-Guanilo, 2017).

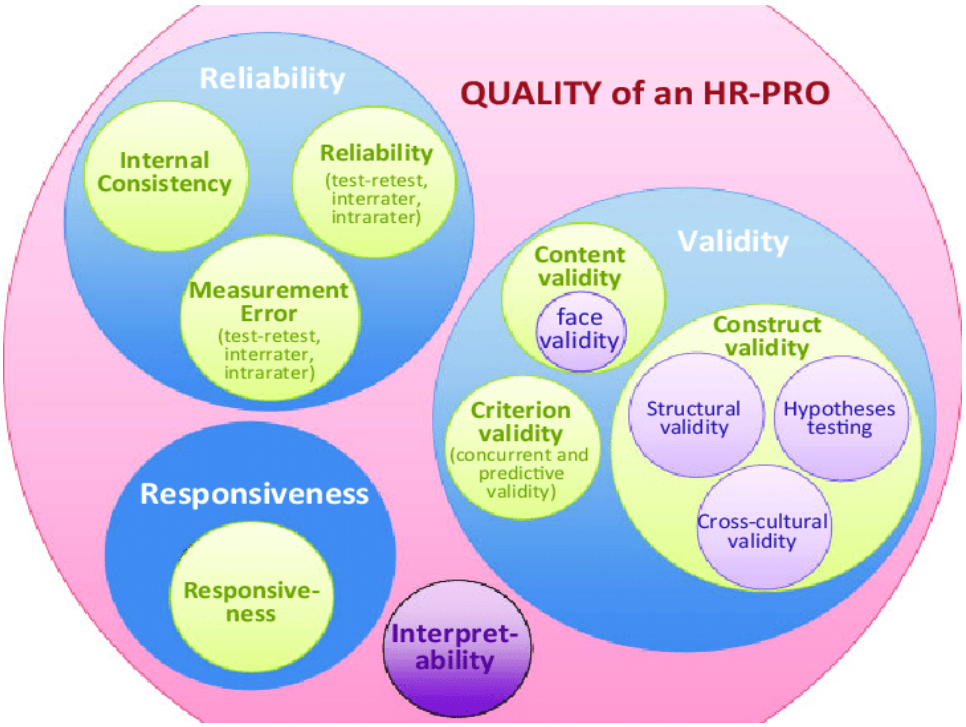

Figura 1 – Propriedades de medidas Fonte: Mokkink, 2016.

Objetivos

Geral: Identificar os instrumentos validados para as Distrofias Musculares categorizando de acordo com os elementos da CIF.

Específicos:

- ✓ Selecionar instrumentos validados para as DM;
- ✓ Analisar a qualidade metodológica dos instrumentos validados selecionados;
- ✓ Analisar se os instrumentos selecionados abordam mais de um domínio da CIF;
- ✓ Propor um modelo de avaliação fisioterapêutica baseado na CIF a partir dos instrumentos encontrados;
- ✓ Dentre os instrumentos selecionados, identificar quais são validados para a população brasileira.

Referências

Adorne, E. D. F., & Bodanese, L. C. (2016). Scientia Medica. *Scientia Medica*, *26*(2), 2–9.

Araújo, E. S. de, & Buchalla, C. M. (2013). Using the ICF in work-related physiotherapy: a contribution to data collection about functioning. *Acta Fisiátrica*, *20*(1), 1–7. https://doi.org/10.5935/0104-7795.20130001

Araujo, E. S., & Buchalla, C. M. (2017). The use of the International Classification of Functioning, Disability and Health in health surveys: a reflexion on its limits and possibilities. *Revista Brasileira de Epidemiologia*, *18*(3), 720–724. https://doi.org/10.1590/1980-5497201500030017

Barra, T. M. F., & Baraldi, K. F. (2013). O uso das escalas funcionais para avaliação clínica da distrofia muscular de Duchenne. *Revista Neurociencias*, *21*(3), 420–426. https://doi.org/10.4181/RNC.2013.21.794.7p

Bos, I., Stallinga, H. A., Middel, B., Kuks, J. B. M., Wynia, K., Classification, I., … Set, N. M. D. C. (2013). *Validation ICF doenças musculos esqueleticas.pdf*. *49*(2), 179–187.

Civil, C. (2019). *Presidência da República*. 1–6.

Conway, K. M., Ciafaloni, E., Matthews, D., Westfield, C., James, K., Paramsothy, P., & Romitti, P. A. (2018). Application of the International Classification of Functioning, Disability and Health system to symptoms of the Duchenne and Becker muscular dystrophies. *Disability and Rehabilitation*, *40*(15), 1773–1780. https://doi.org/10.1080/09638288.2017.1312567

Cunha, C. M., Almeida Neto, O. P. de, & Stackfleth, R. S. (2017). Principais métodos de avaliação psicométrica da confiabilidade de instrumentos de medida. *Revista Brasileira Ciências Da Saúde - USCS*, *14*(49). https://doi.org/10.13037/ras.vol14n49.3671

Diniz, G., Lasmar, L., & Giannetti, J. (2010). Doenças neuromusculares e instrumentos úteis na avaliação motora em crianças e adolescentes. *Revista Médica de Minas Gerais*, *20*(4 Supl 3), 12–19. Retrieved from http://rmmg.org/artigo/detalhes/925

Echevarría-Guanilo ME, Gonçalves N, Romanoski PJ. Propriedades psicométricas de instrumentos de medidas: bases conceituais e métodos de avaliação-parte I. Texto & Contexto-Enfermagem. 2017;26(4)

Guimarães, S. É. R., & Bzuneck, J. A. (2008). Propriedades psicométricas de um instrumento para avaliação da motivação de universitários. *Ciências & Cognição*, *13*(1), 101–113.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., … Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration.

Journal of Clinical Epidemiology, *62*(10), e1-34. https://doi.org/10.1016/j.jclinepi.2009.06.006

Mah, J. K., Korngut, L., Fiest, K. M., Dykeman, J., Day, L. J., Pringsheim, T., & Jette, N. (2015). A Systematic Review and Meta-analysis on the Epidemiology of the Muscular Dystrophies. *Canadian Journal of Neurological Sciences*, *43*(1), 163–177. https://doi.org/10.1017/cjn.2015.311

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., … De Vet, H. C. W. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research*, *19*(4), 539–549. https://doi.org/10.1007/s11136-010-9606-8

Norma Farias, & Cassia Maria Buchalla. (2005). A classificação internacional de funcionalidade, incapacidade e saúde da organização mundial da saúde: conceitos, usos e perspectivas. *Revista Brasileira de Epidemiologia*, *8*(2), 187–193. https://doi.org/10.1590/S1415-790X2005000200011

Pereira, S. M., Castro, E. A. de, & Brochado, V. M. (2016). Neuromuscular diseases and neuromuscular-blocking drugs. *Revista Médica de Minas Gerais*, *26*(Supl 1), 52–59. https://doi.org/10.5935/2238-3182.20160009

Pértile, E., Almeida, G. M. F., Schlindwein-Zanini, R., Fernandes, K. C., & Helegda, L. C. (2014). Utilização do Nintendo Wii na Reabilitação da Distrofia Muscular de Duchenne. *Fiep Bulletin*, *84*(Special Edition), 1–7. https://doi.org/10.16887/85.a2.36

Prinsen CA, Mokkink LB, Bouter LM, Alonso J, Patrick DL, De Vet HC, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Quality of Life Research. 2018;27(5):1147-57

Souza, A. C. de, Alexandre, N. M. C., Guirardello, E. de B., Souza, A. C. de, Alexandre, N. M. C., & Guirardello, E. de B. (2017). Propriedades psicométricas na avaliação de instrumentos: avaliação da confiabilidade e da validade. *Epidemiologia e Serviços de Saúde*, *26*(3), 649–659. https://doi.org/10.5123/S1679-49742017000300022

What is the quality of instruments assessing Activity and Participation in people with muscular dystrophy? A systematic review

Keysy KS Andrade[1], Luana A Soares[2], Caik C Macedo[2], Nelcilaine R Bispo[2], Ricardo R Sousa Junior[3], Vinícius C Oliveira[1,2], Hércules R Leite[1,3], Thaís P Gaiad[1,2]

[1]Postgraduate Program of Rehabilitation and Functional Performance, Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM), Brasil

[2]Physical Therapy Department, Faculty of Health Sciences.Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM), Brasil

[3]Postgraduate Program of Rehabilitation Sciences, School of Physical Education, Physiotherapy and Occupational Therapy. Universidade Federal de Minas Gerais (UFMG), Brasil

*Corresponding author:
Thaís Peixoto Gaiad (Orcid ID 0000-0002-4688-330X)
Universidade Federal dos vales do Jequitinhonha e Mucuri. Departamento de Fisioterapia, Campus JK - Diamantina/MG/Brasil. Rodovia MGT 367 - Km 583, nº 5.000 Alto da Jacuba CEP 39100-000. E-mail address: thais.gaiad@ufvjm.edu.br

What is the quality of instruments assessing Activity and Participation in people with muscular dystrophy? A systematic review

Abstract

Aim To investigate the measurement properties of instruments assessing Activity and Participation for people with Muscular Dystrophy (MD).

Methods A systematic review was conducted on MEDLINE, Embase, AMED, DiTA and PsycINFO. We included published studies that investigated measurement properties of outcome measurement instruments (OMIs) assessing Activity/Participation domains of the International Classification of Functioning (ICF) for MD of any type or age. Two reviewers selected studies, extracted data from the included studies and evaluated the included instruments using the Consensus-based Standards for the selection of health status Measurement Instruments (COSMIN) checklist.

Results Searches identified 6675 references and 46 studies were included with 24 different instruments. Validity was assessed in 34 studies, reliability in 44 of the studies and responsiveness on 6 of them. According to the COSMIN checklist risk of bias, 82.1% of the measurement properties were considered as "very good". The OMIs methodological quality was mainly classified as low evidence (32.14%) due to imprecision and indirectness.

Interpretation Few measurement properties were assessed on the existing OMI for MD and none of them had all properties assessed. The most recommended instruments to assess Activity/Participation are FES-DMD-D2, the MDFRS, MD1-Activ^{C} and the MFM.

Key-words: Muscular dystrophies, Outcome measurement, Physical therapy, Scales, Measurement properties.

What this paper adds:

- There are 24 different instruments validated to evaluated activity/participation for MD.
- The most recommended instrument for DMD to assess activity/participation is FES-DMD-D2
- The MDFRS is the best quality ranking instrument for MD in general
- MFM is the generic instrument with the highest number of evaluated properties
- Evidence quality was mainly classified as low due to imprecision and indirectness
- There is a necessity to develop high quality instruments to assess all MD.

Introduction

Neuromuscular disease comprehends a group of hereditary diseases that involve motor neuron, anterior horn cells, peripheral nerves, neuromuscular junction, and/or the muscle. This heterogeneous group includes Muscular Dystrophy (MD) conditions such as Duchenne, Becker, Limb-Girdle and Facioscapulohumeral dystrophy.[1] MD conditions present autosomal dominant, recessive or X-linked inherited etiologies. They have in common musculoskeletal, cardiorespiratory and autonomic impairments, in consequence of progressive muscle weakness.[2, 3] As a result, individuals in this group present several deficiencies and activity limitations such as swallowing and breathing issues, and loss of mobility throughout their lives, impacting their participation and overall quality of life.[4-6] Multidisciplinary treatment principles for individuals with MD includes identifying impairments, promoting functional activities and participation, in order to prevent the functionality losses.[7] Nowadays, there are several treatment strategies and guidelines for this population with positive effects in outcomes aligned with the International Classification of Functioning, Disability and Health (ICF) framework.[8, 9] In order to properly evaluate these outcomes, standardized and specific outcome measures are extremely necessary.

Standardized outcome measures help clinicians to evaluate patients' strengths and limitations. These include standardized observational tests and patient-reported (family-reported) outcome measures instruments (OMI) with interviews and questionnaires. They help clinicians to identify goal, establish an intervention plan and measure the outcomes changes over the time.[3] Furthermore, considering the natural history of individuals with MD, these assessments will also provide relevant information about the disease progression and longitudinal mobility loss.[5] Following ICF framework it is important to use outcome measures that evaluates the components of body functions and structures (*e.g.* pain, cognitive, sensory, cardiovascular, respiratory, digestive, neuromusculo-skeletal and movement-related functions), but also activity components (*e.g.* execution of self-care, activities of daily living, and mobility), and participation components(*e.g.* frequency and involvement in home, school, and community).[5, 6]

The family and patients' complaints are mainly related to limitations on the ability to get up from the floor, go down and upstairs or walking ability since the first mobility losses are perceived.[10] Thus, clinicians need to change the focus when assessing and setting goals for individuals with dystrophies from just trying to fixing impairments to promote activity and participation.[7] Furthermore, the therapists need to consider if the instrument and its

items reflect the construct or outcome of interest; if it represents relevant outcomes for patients and caregivers; if the administration is feasible and easy to use in the intended context and if it is appropriate for usage in the targeted population.[3] To ensure that all these factors were considered, outcome measures must present well established psychometric properties, which includes content/face validity, internal consistency, criterion validity, construct validity, cross-cultural validity, reproducibility (reliability and measurements of error) and responsiveness.[3, 8, 9]

There are several outcome measures available for evaluation of people with MD. Past reviews highlighted the available "body functions and structures" and "quality of life" instruments; however, little emphasis was given for ICF's activity, and scarcely for participation.[5, 11] The appraisal of available instruments in these domains and their measurement properties can provide evidence-based recommendations to assist clinicians to select proper outcome measures when evaluating patients with MD.[12] Therefore, the aims of this systematic review were: 1) to identify the available instruments validated for infants, adolescents and adults with MD and 2) to analyse the level of evidence of these instruments' measurement properties.

METHOD

This systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) and the Consensus-based Standards for the Selection of Health Status Measurement Instruments guidelines (COSMIN)[13] and it was previously registered at PROSPERO (CRD42020173303).

Study search and eligibility criteria

Search was conducted on databases MEDLINE, Embase, AMED, DiTA and PsycINFO from May to August 2020, without language or date restrictions. We included studies that investigated measurement properties of outcome measures for people with muscular dystrophy of any type or age assessing activity and participation according to ICF.[14] We excluded studies, such as: (1) expert opinion; (2) case reports; (3) if only the abstract was available; and (4) instruments assessing quality of life.

Search strategy consisted of two groups of keywords related to: 'muscular dystrophies' and "psychometric properties". Appendix 1 shows an example of search strategy performed in MEDLINE. Hand-searching in reference list of previous studies were also conducted to identify potential studies not identified in our searches.

Study selection

Selection process and included studies are presented in Figure 1. Two independent reviewers (KKSA and LAS) screened titles and abstracts. Duplicates were removed using EndNote software prior to selection process. After this, full-texts of potentially eligible papers were evaluated. Disagreement between examiners was solved using a third examiner (TPG). All reviewers are physical therapists with expertise with MD.

Data extraction

Two independent reviewers (KKSA and LAS) extracted data using a standardized form with the following information: (i) population (type of muscular dystrophy, sample size of each muscular dystrophy type, age and gender) (ii) setting, country, and language of the outcome instrument, outcome measurements and (iii) measurement properties assessed and the results obtained for each measurement.

Methodological quality assessment

Methodological quality of the included studies was assessed using the newly developed COSMIN Risk of Bias checklist.[12, 13] This checklist assesses risk of bias of measurement properties of systematic reviews studies defined as: cross-cultural validity, measurement error, internal consistency, content validity, structural validity, reliability, construct validity, hypothesis-testing, criterion validity and responsiveness. Each measurement property evaluation consists of several criteria using a 4-point scale: “very good”, “adequate”, “doubtful”, and “inadequate”. Final methodological quality score for each measurement property was determined considering the worst score among all items. Methodological evaluation of interpretability properties studies (ceiling and floor effects and minimal clinically important difference) were not included since there is no specific criteria for these properties.[12]

Quality of the measurement properties of the included studies

Measurement properties reported by each included study were assessed using quality criteria proposed in the COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures.[12] Initially, results for the content validity of each study were rated using five criteria for relevance, one for comprehensiveness, and four for comprehensibility. Results for other measurement properties of each individual study were rated using the updated criteria for good measurement properties as “sufficient (+)”, “insufficient (−)”, or “indeterminate (?)”.[12]

Data from all studies that investigated the same instrument were pooled in order to provide summary of evidence for each measurement property, according to Prinsen.[13]

Quality of pooled measurement properties indexes were rated as “sufficient (+)”, “insufficient (−)”, “indeterminate (?)”, or “inconsistent (±)”.

Quality of Evidence

Quality of evidence was assessed using a modified version of the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach to score the evidence level considering the following four domains: (1) methodological quality (considering the overall methodological quality); (2) inconsistency of results across studies (considering overall quality of pooled measurement properties); (3) imprecision (considering the total sample size of the available studies of each instrument/property)/ and (4) indirectness (considering if the studies presented evidence from different populations than the population of interest in the review). The quality of the evidence was graded as “high”, “moderate”, “low”, or “very low evidence” to each measured property of the included instruments. Each version of the included outcome measure was considered separately in this process.

All these scoring and grading processes were done by two independent reviewers (KKAS e LAS) and a third one was consulted if a consensus was not reached (RRSJ).

RESULTS

Search strategy identified 6.675 publications. After title and abstract selection, 315 publications were potentially eligible and full text was retrieved, 46 of which were included. Most articles were excluded because only the abstract was available and instruments did not comprise activity and participation domains. Details of the selection process are described in Figure 1.

The 46 included studies investigated 24 different OMI for people with MD. Table 1 describes included OMI, its description, scoring criteria, domain/items, administration time. These included instruments presented 84 measurement properties assessed, where 47.61% analysed reliability, 21.4% internal consistency, 9.52% cross-cultural validity, 5.6% responsiveness, 4.76% structural validity, 3.57% measurement error, and 2.38% criterion, face and content validity.

Two groups of outcome measures of Activity and Participation were found: specific and generic. Specific instruments are disease-specific instruments, that is, were designed exclusively to one specific type of MD (*e.g.,*Duchenne muscular dystrophy, Myotonic muscular dystrophy or Facioscapulohumeral muscular dystrophy). Generic-instruments are those designed to any population or those designed to be applied in patients with any

neuromuscular disease, for example the Motor Function Measure (MFM).

Recommendation of the Outcome measurement instruments (OMI) according to GRADE: quality of evidence

The quality of evidence shows that only 14.28% of the properties were classified as high, 28.54% as very low, 32.14% were classified as low evidence, 25% as moderate. Figure 2 shows a decision clinical map based on quality of evidence of the analysed instruments. Figure 2A illustrates the decision framework of 14 specific OMI for Activity assessment.FES-DMD-D2 and DMD-SAT presented better recommendations. Figure 2B illustrates decision framework of 10 generic OMI for Activity assessment.MFM presented the best recommendation among them.

Quality of the Measurement properties

According to the COSMIN checklist risk of bias, from the 46 included studies that analysed instrument's measurement properties for patients with MD, 82.1% were considered as "very good", 25% were considered "adequate", 14.2% were considered "inadequate" and 11.9% were considered "doubtful". Table 2 shows methodological and criteria quality results for all analysed measurement properties.

Measurement properties of the included outcome measurement instruments

The analysis of the 84 measurement properties quality shows that 9.52% were classified as sufficient (+), 9.52% as undetermined (?) and 5.95% as inconsistency (±). For more details about each analysed measurement properties see Table 2.

Reliability

Reliability was assessed on thirty-one studies.[15-43] Quality of evidence for reliability was "high" only for 'Get up and sit on the floor' (FES-DMD-D2) instrument. Instruments which had evidence for reliability classified as "moderate" were Patient-reported outcome measure for upper limb function in Duchenne muscular dystrophy (DMD Upper Limb PROM), Motor Function Measure (MFM), Assessment of the walking activity (FES-DMD-D4), DM1-Activ, DM1-Activ for clinical use (DM1-Activ$^{C)}$, and Myotonic Dystrophy Health Index (MDHI). Reliability evidence was "low" for: North Star Ambulatory Assessment (NSAA), Go up and sit down on a chair (FES-DMD-D1), Assessment of Life Habits (LIFE-H), Dystrophy-specific functional rating scale (MDFRS), Muscular Dystrophy Spine Questionnaire (MDSQ) and Participation and Environment Measure for Children and Youth

(PEM-CY). Ten OMI had “very low” reliability evidence: ULFAS, Motor function measure for Fukuyama congenital muscular dystrophy (GMFM for FSHD), Short Version MFM (MFM-20), Children's Assessment of Participation and Enjoyment and Preferences for Activities of Children (CAPE/PAC), Egen Klassifikation Scale (EK), Egen Klassifikation Scale 2 (EK2), Go up and down stairs (FES-DMD-D3), Neuromuscular-Score (NMS), FSHD composite outcome measure (FSHD-COM) and the Neurological disability score (NDS).[16-18, 27, 28, 31, 33-35, 41, 42, 44-51]

Measurement error

Four studies assessed measurement error property.[15, 52-54] FES-DMD-D4 was presented “moderate” evidence, CAPE/PAC had it classified as “low” and FSHD-COM as “very low”.

Internal consistency

Twenty-two studies assessed internal consistency.[15-20, 24, 27, 28, 32-36, 38, 39, 42, 43, 50, 55-57] DM1-Activ^{C}, Functional Ability Self-Assessment Tool (DMD SAT), MDHI, and MDFRS presented high quality of evidence. Internal Consistency of MFM was classified as “moderate” and another eight OMI had “low” evidence: DMD Upper Limb PROM, NSAA, CAPE/PAC, Community Integration Questionnaire (CIQ), DM1-Activ, FSHD-COM, NDS and PEM-CY. Five instruments presented “very low” quality of evidence for internal consistency: ACTIVLIM questionnaire, ULFAS, MFM-20, EK and EK2.[15-19, 27, 29-31, 33-36, 38, 41-44, 50, 51, 54-56, 58-64]

Hypotheses Testing

Twenty-three studies assessed Hypotheses testing for construct validity property. [16-18, 24, 25, 27, 28, 31-39, 42, 43, 52, 55, 62, 63, 65]. It was classified as “high” for DMD Upper Limb PROM, DM1-Activ, DM1-Activ^{C} and MDFRS. The FES-DMD-D4, GMFM for FCMD, MFM, PEM-CY, MDHI, ABILHAND and CIQ presented “moderate” evidence. NSAA, FSHD-COM, MDSQ, NDS, and MFM-20 presented “low evidence”. Only NM-Score had “very low” evidence for hypothesis testing.

Cross-cultural validity

Fourteen studies assessed cross-cultural validity.[15, 19, 30, 31, 34, 35, 42, 43, 46, 50, 51, 59, 63, 64] NM-Score and the MDHI had this property classified as “low” and another six OMI had it

classified as “very low”: NSAA, EK, EK2, MFM, PEM-CY and CAPE/PAC.

Structural validity

Five studies assessed structural validity.[36, 40, 44, 46, 55] Only MDFRS had quality of evidence classified as “high” and three ones presented moderate evidence: MFM, NSAA and PEM-CY.

Content/Face validity

Only two studies investigated content validity.[46, 62] ABILHAND questionnaire had quality of evidence for content validity classified as “low” and NMS classified as “very low”.[38, 62]

Criterion validity

Two studies assessed criterion validity.[46, 66] Quality of evidence was moderate for MFM and “low” for NMS.

Responsiveness

Seven studies assessed responsiveness property.[28, 35, 53, 62, 67, 68] MDFRS, MFM and ACTIVLIM questionnaire presented moderate quality of evidence. FES-DMD-D1, FES-DMD-D3 and FES-DMD-D4 presented low quality of evidence.[36]

DISCUSSION

This systematic review collected 24 instruments assessing activity/participation of people with MD which have mainly ‘low’ recommendation according to their measurement properties already investigated. DMD is the population with more available specific instruments, followed by myotonic MD and other instruments to the general population of muscular dystrophies. The three OMI with high quality of evidence are instruments specific for DMD population: FES-DMD, DMD SAT and DMD Upper limb.

Quality of the evidence of instrument’s measurement properties refers to the degree of instruments’ available trustworthiness. This enables reviewers to draw transparent conclusions and make recommendations on the quality of outcome measures, and supports evidence-based selection for use in research and in clinical practice.[19] The measurement properties more frequently investigated were the reliability through test-retest analysis,

internal consistency and construct validity through hypothesis testing.

Test-retest reliability of D2 domain of the FES-DMD has “high” evidence which means that it is highly recommended to be used at clinical practice with few random errors.[69] FES-DMD is an observed-rater functional scale specific for activity assessment that encompasses four domains: sitting and standing from a chair (D1), going up and down stairs (D3), sitting and standing from the ground (D2) and walking (D4). Fernandes et al[69] investigated the reliability of FES-DMD-D3 stratifying by age, Vignos scale and time to perform activities. They found a moderate to weak correlation of these variables in the studied population but our analysis using the COSMIN attests that FES DMD-D3 has a “very low” reliability. It was downgraded due to the sample size lower than 100 and risk of bias of one study classified as ‘adequate’. The D1 and D4 FES-DMD domains had the test-retest reliability also assessed but they were classified as low and moderate, respectively.

Internal consistency of DMD SAT instrument was classified as “high” showing that the items of this OMI are homogenous, measuring the same construct and producing consistent results. It is a patient-reported instrument that assesses functional capacity of ambulation for people with DMD. It is used at clinical or research settings and can be answered by patient or its caregiver without the presence of a health professional. It presents 4 domains: arm function, mobility, transfer and ventilation status[58] and has its time of administration not reported.

Another recommended OMI is the DMD Upper Limb PROM, a patient-reported instrument that had its hypotheses testing to investigate construct validity classified as “high” which means that in its totality this OMI really measures the performance of the upper limb on activities of daily life of people with DMD through patient-report. It has 32 items and four domains of activity of daily life: feeding, self-care, home environment, taking 10 minutes for administration. The study of Klingers et al[70] shows that the test-retest reliability has an excellent result of 0.99 reflecting the degree of consistency of the instrument items.[71] According to Davoli et al[72] Perfomance of the Upper Limb (PUL) and DMD Upper Limb PROM scale have the best methodological quality, based on scientific evidence used to assess upper-limb function in patients with DMD. PUL is an observed–rater instrument designed to people with DMD but it was not included in our analysis because it does not assess Activity/Participation. Besides, following the COSMIN risk of bias and checklist, our analysis also shows that the Upper Limb PROM is one of the three OMI with the best recommendation, with only three, but important psychometric properties already tested: hypothesis test, test-retest reliability and internal consistency tested but not all of them are

classified as high.

Based on the Hammersmith Functional Motor Scale, the NSAA is an observer-rater instrument disease-specific that assesses mobility through 17 items related to functional capacity of boys with DMD.[41] Mazzone et al[41] attested a good reliability intra- and inter-observer and another study Eagle et al[73] attested the reliability and viability of detection of change over time. Our data show that the NSAA has a moderate level of evidence to structural validity being downgraded at the item risk of bias. The sample to attest internal consistency and hypothesis testing was lower than 100 and included other health conditions that were not MD which motivated the OMI to be downgraded. Cross-cultural validity was also downgraded due to the sample size (lower than 100), other health conditions and one study considered "inadequate" which led to a final GRADE of "very low" to this property. Okama et al[42] applied the Portuguese version of the NSAA in 89 patients with DMD aging 4 to 17 years old to assess the internal consistency which was considered good (alpha Cronbach of 0,70 a 0,95). In this way, both NSAA and FES-DMD are designed to assess people with DMD with the NSAA instrument presenting a 'moderate' recommendation and the FES-DMD-D2 presenting a 'high' recommendation of use.

DM1-Activ^{C} is specific for myotonic MD and is a well recommended OMI. It presented high quality of evidence for test-retest reliability, internal consistency and hypotheses test. This OMI was designed for clinical use, presenting 25 items assessing Activity. According to the study of Hermans et al[16] the DM1-Activ^{C} encompass the Rash model and presents strong validity and reliability which means that is has a positive ability to distinguish between different degrees of disability. The DM1-Activ^{C} differs from the DM1-Activ due its amount of items that are 49 representing social and daily activities and three options of answer totalizing a maximum of five points for item: 0 = enable to perform it; 1 = able to perform but with difficult; 2 = easy to perform it. This last one is not so well recommended as DM1-Activ^{C} because reviewers have argued that its validity had included only 186 patients which affected the model robustness, the definition of some items was vague, the answers ambiguous that could lead to mistakes at its interpretation. MDFRS is an instrument that assesses the global impact and severity of MD.[35] Quality of evidence for reliability was classified as "low" due to the small sample size and only one study with adequate quality.

The majority of the instruments for people with MD have been downgraded at the GRADE recommendation at the 'imprecision' item which analyses the sample size. When the study used a sample size of less than 100 it was downgraded at 1 point and when the sample

size was lower than 50 the instrument was downgraded 2 points.[12] Another limitation of the included studies was indirectness. They have downgraded in the quality of evidence analysis due to the inclusion of other health conditions at the sample, leading to a downgrade of 1 point in the majority of the instruments. Other types of neuromuscular disorders were included in the studies of the psychometric properties besides the population of our interest. MD is a rare condition which does not enable researchers to adequately compose intervention studies with a satisfactory sample size to reach precise conclusions. It reflects the limitation pointed out in the last systematic reviews about the exercise on MD conducted by Voet et al[74] and Gianolla et al[75] of inconclusive results due to the small sample size (imprecision) and variability (indirectness). Future studies might evaluate and suggest quality of evidence criteria for instruments designed for rare health-conditions.

None of the included generic OMI received a high quality of evidence but five of them presented moderate levels of evidence: MFM for Structural validity, Internal consistency, Cross-cultural validity, Reliability, Validity Criterion, Hypotheses testing and responsiveness, PEM-CY for Internal consistency, Cross-cultural validity, Reliability and Hypotheses testing, ABILHAND for Content validity and Hypotheses Testing, ACTIVLIM for Internal Consistency and Responsiveness and CIQ for structural validity. MFM is an observer-rater OMI that has the highest number of assessed psychometric properties when compared to all included OMI and the best methodological quality among generic instruments. It is the instrument most used at the clinical practice with people with MD in general,[66] but was designed to assesses the severity and motor function progression and rehabilitation outcomes for people with neuromuscular disorders through three domains: Standing position and transferring; axial and proximal motor function; distal motor function.[30] Internal consistency and reliability presents a "moderate" level of evidence, being downgraded at the 'indirectness' item due to including different health conditions in its GRADE analysis. According to the studied properties, the MFM can be reliably applied on MD and other neuromuscular conditions.

The responsiveness was assessed only on six instruments for people with MD: MDFRS, MFM, ACTIVLIM questionnaire (moderate) and FES-DMD-D1, FES-DMD-D3 and the FES-DMD-D4 (low). This property, also called sensibility, says about the instrument structure and the detection of changes at the assessed construct.[71] Responsiveness is considered an important longitudinal construct property with many included studies not considering it. Besides the paucity of studies about responsiveness detected in our data, the knowledge of this property allows the detection of clear intervals of assessment to each

instrument/health condition. In the study of Vuillerot et al,[45] they attested that one year is a short time to detect changes in neuromuscular disorders of slow progression. It is a crucial measurement property once it will attest with security that the selected instrument will report the changes over time of a progressive health condition with different speeds of evolution as muscular dystrophy conditions.

Strengths and limitations

This work is the first to systematically synthesize the methodological quality of the outcome instruments designed to assess Activity and Participation of pwMD according to the recommendations of the COSMIN Risk of Bias checklist. To each one of the included instruments, the modes and time of administration, items and domains and target subtypes of muscular dystrophy population were detailed. The analysis of the methodological quality enables the recommendation of each one of the outcome instruments on the light of the COSMIN.

The included instruments encompasses the Activity/Participation component of ICF although this systematic review do not aims to delimit which or how much domains of activity or participation are assessed in each one of the instruments. It is a linking study that needs further investigation in future research. The low sample size of the psychometrics studies and some measures not adopted by the COSMIN checklist hampered the analysis of some outcome instruments that had their evidence level downgraded. The absence of outcome instruments designed for different types of muscular dystrophy do not enable the recommendation of high quality instruments for the muscular dystrophy population and its subtypes.

CONCLUSION

There are 24 instruments of activity/participation assessment of pwMD with its measurement properties already attested with test-retest reliability, internal consistency and hypothesis testing the most frequently assessed. The six best recommended instruments for this population are the FES-DMD-D2, DMD SAT, DMD-UL-PROM, DM1-Activ-C, DM1-Activ and the MDFRS. None of the instruments for people with MD had all its properties assessed and the majority of the instruments with the best recommendation are designed for the DMD population, followed by myotonic MD. This study reveals a need for future measurement properties analysis of instruments for all subtypes of MD and following

COSMIN guidelines, promoting adequate assessment for researchers and clinicians and supporting MD progression assessment.

REFERENCES

1. Bhatt JM. The Epidemiology of Neuromuscular Diseases. Neurologic Clinics. 2016;34(4):999-1021.
2. Jorge MGAV, Carrapatoso BC, Fernandes ABS. A Fisioterapia na Amiotrofia Espinhal Progressiva Tipo I. Revista Neurociências. 2013;21(3):402-7.
3. Wright V, Majnemer A, Maltais DB, Burtner PA, Sanders H, editors. Motor measures: A moving target? Seminars in pediatric neurology; 2013: Elsevier.
4. Palisano RJ, Campbell SK, Orlin M. Physical Therapy for Children-E-Book: Elsevier Health Sciences; 2014.
5. Diniz GPC, LMLBF L, JG G. Doenças neuromusculares e instrumentos úteis na avaliação motora em crianças e adolescentes. Rev Med Minas Gerais. 2010;20(4):12-9.
6. Conway KM, Ciafaloni E, Matthews D, Westfield C, James K, Paramsothy P, et al. Application of the International Classification of Functioning, Disability and Health system to symptoms of the Duchenne and Becker muscular dystrophies. Disability and rehabilitation. 2018;40(15):1773-80.
7. Bendixen RM, Lott DJ, Senesac C, Mathur S, Vandenborne K. Participation in daily life activities and its relationship to strength and functional measures in boys with Duchenne muscular dystrophy. Disability and rehabilitation. 2014;36(22):1918-23.
8. Souza ACd, Alexandre NMC, Guirardello EdB. Propriedades psicométricas na avaliação de instrumentos: avaliação da confiabilidade e da validade. Epidemiologia e Serviços de Saúde. 2017;26:649-59.
9. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. Journal of clinical epidemiology. 2007;60(1):34-42.
10. Jerath NU, Simoens K, Mann D, Kollasch S, Grosland N, Malik KA, et al. Survey of the functional priorities in patients with disability due to neuromuscular disorders. Disability and Rehabilitation: Assistive Technology. 2019;14(2):133-7.
11. Powell PA, Carlton J, Rowen D, Brazier JE. Measuring Quality of Life for Economic Evaluation in Duchenne Muscular Dystrophy (Dmd) from Childhood to Adulthood: Protocol for a New Preference-Based Measure. Value in Health. 2018;21 (Supplement 3):S473.
12. Mokkink LB, De Vet HC, Prinsen CA, Patrick DL, Alonso J, Bouter LM, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. Quality of Life Research. 2018;27(5):1171-9.
13. Prinsen CA, Mokkink LB, Bouter LM, Alonso J, Patrick DL, De Vet HC, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Quality of Life Research. 2018;27(5):1147-57.
14. Organization WH. IFC: International Classification of Functioning, Disability and Health. 2001.
15. Fink A, Gebhard B, Erdwiens S, Haddenhorst L, Nowak S. Reliability of the German version of the Children's Assessment of Participation and Enjoyment (CAPE) and Preferences for Activities of Children (PAC). Child: care, health and development. 2016;42(5):683-91.
16. Hermans MC, Faber CG, De Baets MH, de Die-Smulders CE, Merkies IS. Rasch-built myotonic dystrophy type 1 activity and participation scale (DM1-Activ). Neuromuscular Disorders. 2010;20(5):310-8.

17. Hermans MC, Hoeijmakers JG, Faber CG, Merkies IS. Reconstructing the Rasch-Built Myotonic Dystrophy Type 1 Activity and Participation Scale. PLoS ONE [Electronic Resource]. 2015;10(10):e0139944.
18. Klingels K, Mayhew A, Mazzone ES, Eagle M, Duong T, Decostre V, et al. Development of a patientreported outcome measure for upper limb function in duchenne muscular dystrophy (DMD-upper limb prom). Journal of Neuromuscular Diseases. 2016;3 (Supplement 1):S140-S1.
19. Fagoaga J, Girabent-Farres M, Bagur-Calafat C, Febrer A, Steffensen BF. Translation and validation of the Egen Klassifikation scale for the Spanish population: Functional assessment for non-ambulatory individuals with Duchenne's muscular dystrophy and spinal muscular atrophy. Revista de Neurologia. 2013;56(11):555-61.
20. Escobar RG, Lucero N, Solares C, Espinoza V, Moscoso O, Olguin P, et al. [Upper limb functional assessment scale for children with Duchenne muscular dystrophy and Spinal muscular atrophy]. Revista Chilena de Pediatria. 2016;16:16.
21. Hukuda ME, Escorcio R, Fernandes LA, de Carvalho EV, Caromano FA. Evaluation scale development, reliability for sitting and standing from the chair for Duchenne muscular dystrophy. Journal of motor behavior. 2013;45(2):117-26.
22. Escorcio R, Caromano FA, Hukuda ME, Fernandes LA. Development of an evaluation scale for sitting and standing from the ground for children with Duchenne muscular dystrophy. Journal of motor behavior. 2011;43(1):31-6.
23. de Carvalho EV, Hukuda ME, Escorcio R, Voos MC, Caromano FA. Development and Reliability of the Functional Evaluation Scale for Duchenne Muscular Dystrophy, Gait Domain: A Pilot Study. Physiotherapy research international : the journal for researchers and clinicians in physical therapy. 2015;20(3):135-46.
24. Eichinger K, Currence M, Capron K, Duong T, Gee R, Herbelin L, et al. Muscle strength and function measures in a multicenter study of myotonic dystrophy type 1 (DM1): Baseline impairment and test-retest agreement over 3 months. Neurology Conference: 70th Annual Meeting of the American Academy of Neurology, AAN. 2018;90(15 Supplement 1).
25. Sato T, Adachi M, Nakamura K, Zushi M, Goto K, Murakami T, et al. The gross motor function measure is valid for Fukuyama congenital muscular dystrophy. Neuromuscular Disorders. 2017;27(1):45-9.
26. Gagnon C, Mathieu J, Noreau L. Measurement of participation in myotonic dystrophy: reliability of the LIFE-H. Neuromuscular Disorders. 2006;16(4):262-8.
27. Berard C, Payan C, Hodgkinson I, Fermanian J. A motor function measure scale for neuromuscular diseases. Construction and validation study. Neuromuscular Disorders. 2005;15(7):463-70.
28. Berard C, Payan C, Fermanian J, Girardot F. A motor function measurement scale for neuromuscular diseases - Description and validation study. [French]. Revue Neurologique. 2006;162(4):485-93.
29. Ruiz-Cortes X, Ortiz-Corredor F, Mendoza-Pulido C. Reliability of home-based, motor function measure in hereditary neuromuscular diseases. Journal of International Medical Research. 2017;45(1):261-71.
30. Iwabe C, Miranda-Pfeilsticker B, Nucci A. Medida da função motora: versão da escala para o português e estudo de confiabilidade. Brazilian Journal of Physical Therapy. 2008;12(5):417-24.
31. Huang J, Tepelus TC, Baghdasaryan E, Huang P, Shi Y, Hsu HY, et al. Correlation between Guttata Severity and Thickness of Descemet's Membrane and the Central Cornea. Current Eye Research. 2019;44(8):849-55.
32. De Lattre C, Payan C, Berard C. Motor function measure: Construction of a short form (MFM-20) for children with neuromuscular disease aged between 2 and 6. Annals of Physical

and Rehabilitation Medicine. 2011;1):e199-e200.
33. Heatwole C, Bode R, Johnson N, Dekdebrun J, Dilek N, Heatwole M, et al. Myotonic Dystrophy Health Index: Initial evaluation of a disease-specific outcome measure. Muscle and Nerve. 2014;49(6):906-14.
34. Sansone VA, Lizio A, Greco L, Gragnano G, Zanolini A, Gualandris M, et al. The Myotonic Dystrophy Health Index: Italian validation of a disease-specific outcome measure. Neuromuscular Disorders. 2017;27(11):1047-53.
35. Mori I, Fujino H, Matsumura T, Takada H, Ogata K, Nakamori M, et al. The myotonic dystrophy health index: Japanese adaption and validity testing. Muscle and Nerve. 2019;59(5):577-82.
36. Lue Y, Su C, Yang R, Su W, Lu Y, Lin R, et al. Development and validation of a muscular dystrophy-specific functional rating scale. Clinical Rehabilitation. 2006;20(9):804-17.
37. Wright JG, Smith PL, Owen JL, Fehlings D. Assessing functional outcomes of children with muscular dystrophy and scoliosis: The muscular dystrophy spine questionnaire. Journal of Pediatric Orthopaedics. 2008;28(8):840-5.
38. Contardi S, Pizza F, Falzone F, D'Alessandro R, Avoni P, Di Stasi V, et al. Development of a disability scale for myotonic dystrophy type 1. Acta Neurologica Scandinavica. 2012;125(6):431-8.
39. Vuillerot C, Rippert P, Kinet V, Renders A, Jain M, Waite M, et al. Rasch analysis of the motor function measure in patients with congenital muscle dystrophy and congenital myopathy. Annals of Physical and Rehabilitation Medicine. 2014;1):e97.
40. Mayhew AG, Coratti G, Mazzone ES, Klingels K, James M, Pane M, et al. Performance of Upper Limb module for Duchenne muscular dystrophy. Developmental Medicine & Child Neurology. 2020;62(5):633-9.
41. Mazzone ES, Messina S, Vasco G, Main M, Eagle M, D'Amico A, et al. Reliability of the North Star Ambulatory Assessment in a multicentric setting. Neuromuscular Disorders. 2009;19(7):458-61.
42. Okama LO, Zampieri LM, Ramos CL, Toledo FO, Alves CRJ, Mattiello-Sverzut AC, et al. Reliability and validity analyses of the North Star Ambulatory Assessment in Brazilian Portuguese. Neuromuscular Disorders. 2017;27(8):723-9.
43. Kaya Kara O, Turker D, Kara K, Yardimci-Lokmanoglu BN. Psychometric properties of the Turkish version of Participation and Environment Measure for Children and Youth. Child: Care, Health and Development. 2020;46(6):711-22.
44. Guillot T, Roche S, Rippert P, Hamroun D, Iwaz J, Ecochard R, et al. Is Going Beyond Rasch Analysis Necessary to Assess the Construct Validity of a Motor Function Scale? Archives of Physical Medicine and Rehabilitation. 2018;99(9):1776-82.e9.
45. Vuillerot C, Payan C, Girardot F, Fermanian J, Iwaz J, Berard C, et al. Responsiveness of the motor function measure in neuromuscular diseases. Archives of Physical Medicine and Rehabilitation. 2012;93(12):2251-6.e1.
46. Vuillerot C, Meilleur KG, Jain M, Waite M, Wu T, Linton M, et al. English cross-cultural translation and validation of the neuromuscular score: a system for motor function classification in patients with neuromuscular diseases. Archives of Physical Medicine & Rehabilitation. 2014;95(11):2064-70.e1.
47. Huang M, Cao J, Sun J, Li W, Qin L, Li H, et al. Cross-cultural Adaptation and Multi-centric Validation of the Motor Function Measure Chinese Version (MFM-32-CN) for Patients with Neuromuscular Diseases. Developmental Neurorehabilitation. 2020;23(4):210-7.
48. Gagnon C, Chouinard MC, Laberge L, Brisson D, Gaudet D, Lavoie M, et al. Prevalence of lifestyle risk factors in myotonic dystrophy type 1. Canadian Journal of

Neurological Sciences. 2013;40(1):42-7.
49. Mayhew A, Cano S, Scott E, Eagle M, Bushby K, Muntoni F, et al. Moving towards meaningful measurement: Rasch analysis of the North Star Ambulatory Assessment in Duchenne muscular dystrophy. Developmental Medicine & Child Neurology. 2011;53(6):535-42.
50. Alemdaroglu I, Karaduman A, Iyigun-Yatar G, Tunca-Yilmaz O, Topaloglu H. Turkish version of the Egen Klassifikation scale version 2: Validity and reliability in the Turkish population. Turkish Journal of Pediatrics. 2014;56(6):643-50.
51. Fagoaga J, Girabent-Farres M, Bagur-Calafat C, Febrer A, Steffensen BF. Functional assessment for people unable to walk due to spinal muscular atrophy and Duchenne muscular dystrophy. Translation and validation of the Egen Klassifikation 2 scale for the Spanish population. [Spanish]. Revista de Neurologia. 2015;60(10):439-46.
52. De Carvalho E, Hukuda M, Voos M, Caromano F. Responsiveness of the functional evaluation scale for patients with Duchenne muscular dystrophy-domain 2: Gait. Neuromuscular Disorders. 2015;2):S231.
53. de Carvalho EV, Caromano FA, Goya PSA, Hukuda ME, Voos MC. Responsiveness of the Gait Domain of the Functional Evaluation Scale for Children with Duchenne Muscular Dystrophy. Physical & occupational therapy in pediatrics. 2019;39(1):107-18.
54. Eichinger K, Heatwole C, Iyadurai S, King W, Baker L, Heininger S, et al. Facioscapulohumeral muscular dystrophy functional composite outcome measure. Muscle and Nerve. 2018;58(1):72-8.
55. Hirsh AT, Braden AL, Craggs JG, Jensen MP. Psychometric properties of the Community Integration Questionnaire in a heterogeneous sample of adults with physical disability. Archives of Physical Medicine and Rehabilitation. 2011;92(10):1602-10.
56. Batcho CS, Roy AJ, Thonnard JL, Penta M, de Noordhout AM, Servais L, et al. How robust is ACTIVLIM for the follow-up of activity limitations in patients with neuromuscular diseases? Neuromuscular Disorders. 2016;26(3):211-20.
57. Landfeldt E, Lindgren P, Bell CF, Schmitt C, Guglieri M, Straub V, et al. Compliance to care guidelines for duchenne muscular dystrophy. Journal of Neuromuscular Diseases. 2015;2(1):63-72.
58. Landfeldt E, Mayhew A, Eagle M, Lindgren P, Bell C, Guglieri M, et al. Development and psychometric analysis of the Duchenne muscular dystrophy Functional Ability Self-Assessment Tool (DMDSAT). Neuromuscular Disorders. 2015;25(12):937-44.
59. Gagnon C, Tremblay M, CoTe I, Heatwole C. French translation and cross-cultural adaptation of The Myotonic Dystrophy Health Index. Muscle and Nerve. 2018;57(4):686-9.
60. Benaïm C, Sacconi S, Fournier-Mehouas M, Tanant V, Desnuelle C. Analyse de validité de la «Mesure de la fonction motrice»(MFM) en pratique de consultation adulte d'un centre de référence pour maladies neuromusculaires. Revue neurologique. 2010;166(1):49-53.
61. Mayhew AG, Eagle M, Willis TA, Straub V. Exploratory Rasch analysis of adapted North Star ambulatory assessment in LGMD 2I. Neuromuscular Disorders. 2011;21 (9-10):667.
62. Vandervelde L, Van Den Bergh PYK, Penta M, Thonnard JL. Validation of the ABILHAND questionnaire to measure manual ability in children and adults with neuromuscular disorders. Journal of Neurology, Neurosurgery and Psychiatry. 2010;81(5):506-12.
63. Martinez JA, Brunherotti MA, de Assis MR, Sobreira CF. [Validation of the EK functional motor scale in the Portuguese language]. Revista Da Associacao Medica Brasileira. 2006;52(5):347-51.
64. Steffensen B, Hyde S, Lyager S, Mattsson E. Validity of the EK scale: a functional assessment of non-ambulatory individuals with Duchenne muscular dystrophy or spinal

muscular atrophy. Physiotherapy Research International. 2001;6(3):119-34.
65. Sato T, Adachi M, Matsuo A, Zushi M, Goto K, Hirose M, et al. A short form of gross motor function measure for Fukuyama congenital muscular dystrophy. Brain and Development. 2020;42(5):383-8.
66. Benaim C, Sacconi S, Fournier-Mehouas M, Tanant V, Desnuelle C. Validity of the motor function measurement scale when routinely used in the follow-up of adult outpatients in a neuromuscular center. [French]. Revue Neurologique. 2010;166(1):49-53.
67. Hukuda ME, Caromano FA, Escorcio R, Carvalho EV, Blascovi-Assis SM, Voos MC. How often should sitting and rising from a chair be evaluated in patients with Duchenne muscular dystrophy? Arquivos de Neuro-Psiquiatria. 2017;75(9):625-30.
68. Albuquerque PS. Responsividade do domínio subir e descer escada da escala de avaliação funcional para pessoas com distrofia muscular de Duchenne, no período de um ano. 2016. p. [50]-[].
69. Fernandes LA, Caromano FA, Assis S, Hukuda ME, Voos MC, Carvalho EV. Relação entre escore FES-DMD-subir e descer escada com escore Escala Vignos, idade e tempo de realização das atividades em meninos com Distrofia Muscular de Duchenne. Brazilian Journal of Physical Therapy. 2014;18(6):513-20.
70. Klingels K, Mayhew A, Mazzone E, Main M, Decostre V, Van Den Hauwe M, et al. Development of a patient-reported outcome measure for arm and hand function in Duchenne muscular dystrophy (UL-PROM DMD). Neuromuscular Disorders. 2015;2):S227.
71. Echevarría-Guanilo ME, Gonçalves N, Romanoski PJ. Propriedades psicométricas de instrumentos de medidas: bases conceituais e métodos de avaliação-parte I. Texto & Contexto-Enfermagem. 2017;26(4).
72. Davoli GBDQ, Cardoso J, Silva GC, Moreira RdFC, Mattiello-Sverzut AC. Instruments to assess upper-limb function in children and adolescents with neuromuscular diseases: a systematic review. Developmental Medicine & Child Neurology. 2021.
73. McDonald CM, Henricson EK, Abresch RT, Duong T, Joyce NC, Hu F, et al. Long-term effects of glucocorticoids on function, quality of life, and survival in patients with Duchenne muscular dystrophy: a prospective cohort study. The Lancet. 2018;391(10119):451-61.
74. Voet NB, van der Kooi EL, van Engelen BG, Geurts AC. Strength training and aerobic exercise training for muscle disease. Cochrane Database of Systematic Reviews. 2019(12).
75. Clutterbuck GL, Auld ML, Johnston LM. High-level motor skills assessment for ambulant children with cerebral palsy: A systematic review and decision tree. Developmental Medicine & Child Neurology. 2020;62(6):693-9.
76. Vandervelde L, Van den Bergh PYK, Goemans N, Thonnard JL. Activity limitations in patients with neuromuscular disorders: A responsiveness study of the ACTIVLIM questionnaire. Neuromuscular Disorders. 2009;19(2):99-103.
77. Hukuda ME, Caromano FA, Escorcio R, De Carvalho EV, Blascovi-Assis SM, Voos MC. How often should sitting and rising from a chair be evaluated in patients with Duchenne muscular dystrophy? Arquivos de Neuro-Psiquiatria. 2017;75(9):625-30.
78. Fernandes LAY, Caromano FA, Hukuda ME, Escorcio R, Carvalho EV. Elaboration and reliability of functional evaluation on going up and downstairs scale for Duchenne Muscular Dystrophy. [Portuguese]. Revista Brasileira de Fisioterapia. 2010;14(6):518-26.
79. Eichinger K, Heatwole C, Iyadurai S, King W, Baker L, Heininger S, et al. Facioscapulohumeral muscular dystrophy functional composite outcome measure. Muscle Nerve. 2018.
80. Gagnon C, Mathieu J, Noreau L. Measurement of participation in myotonic dystrophy: Reliability of the LIFE-H. Neuromuscular Disorders. 2006;16(4):262-8.
81. Berard C, Payan C, Fermanian J, Girardot F, Groupe d'Etude MFM. [A motor function

measurement scale for neuromuscular diseases - desc2ription and validation study]. Revue Neurologique. 2006;162(4):485-93.
82. Mori I, Fujino H, Matsumura T, Takada H, Ogata K, Nakamori M, et al. The myotonic dystrophy health index: Japanese adaption and validity testing. Muscle & Nerve. 2019;59(5):577-82.
83. Vuillerot C, Meilleur KG, Jain M, Waite M, Wu T, Linton M, et al. English cross-cultural translation and validation of the neuromuscular score: A system for motor function classification in patients with neuromuscular diseases. Archives of Physical Medicine and Rehabilitation. 2014;95(11):2064-70.
84. Vuillerot C, Rippert P, Jain M, Waite M, Glanzman A, Kinet V, et al. The Rasch-scaled motor function measure for patients with congenital disorders of muscle. Neuromuscular Disorders. 2013;23 (9-10):791-2.
85. Mayhew A, Cano S, Scott E, Eagle M, Bushby K, Muntoni F, et al. Moving towards meaningful measurement: Rasch analysis of the North Star Ambulatory Assessment in Duchenne muscular dystrophy. Developmental Medicine and Child Neurology. 2011;53(6):535-42.
86. Vandervelde L, Van den Bergh PY, Goemans N, Thonnard JL. Activity limitations in patients with neuromuscular disorders: a responsiveness study of the ACTIVLIM questionnaire. Neuromuscular Disorders. 2009;19(2):99-103.
87. Fernandes LA, Caromano FA, Hukuda ME, Escorcio R, Carvalho EV. Elaboration and reliability of functional evaluation on going up and downstairs scale for Duchenne Muscular Dystrophy. Revista Brasileira de Fisioterapia. 2010;14(6):518-26.
88. Vuillerot C, Meilleur K, Jain M, Waite M, Wu T, Datsgir J, et al. English cross-cultural translation and validation of the NM-Score: A system for motor function classification in patients with neuromuscular diseases. Journal of Neuromuscular Diseases. 2014;1 (Supplement 1):S387.

Figure 1 - Flowart

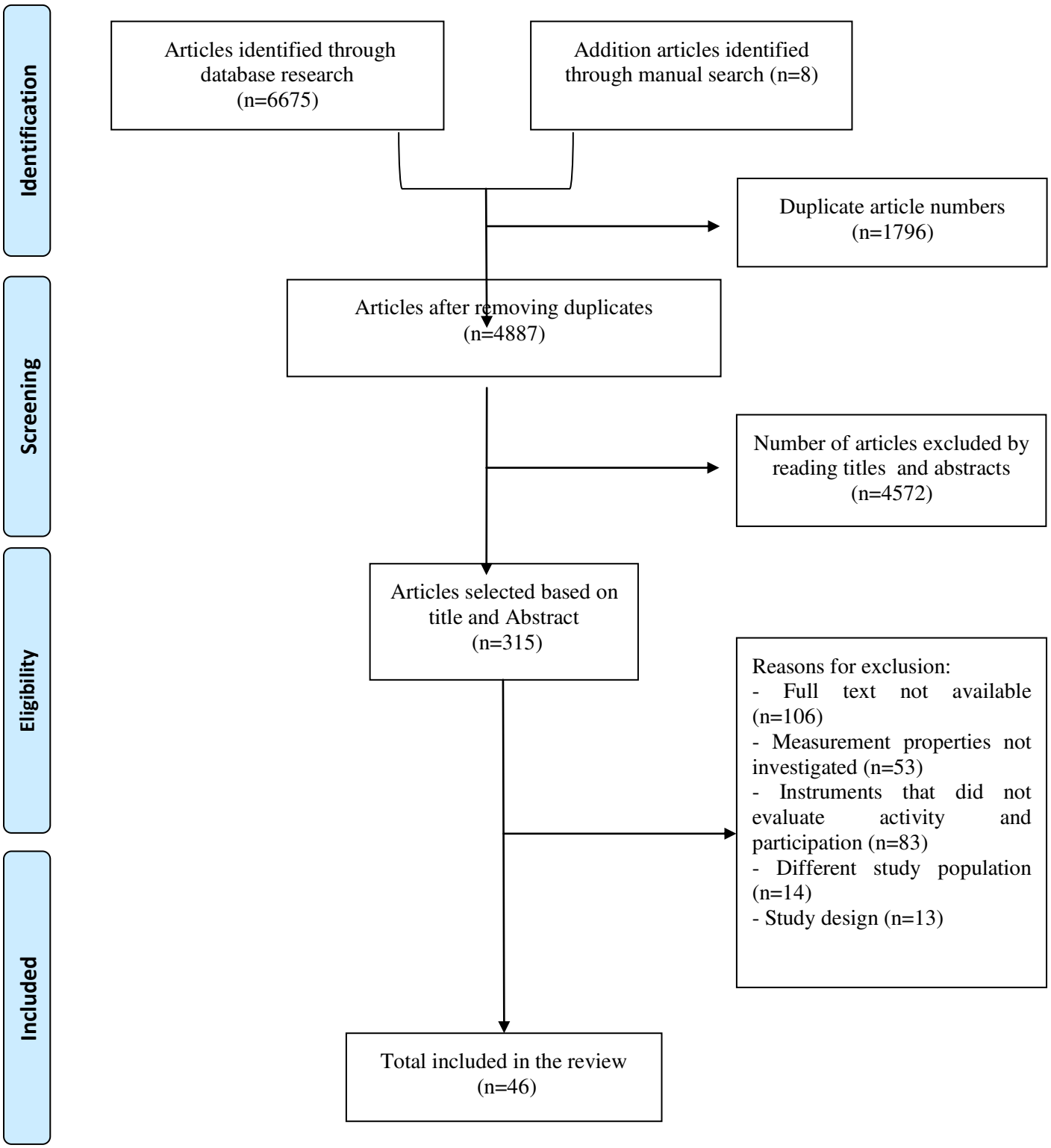

Figure 2A - Decision framework for pwMD-specific Outcome Measurement selection for Activity assessment

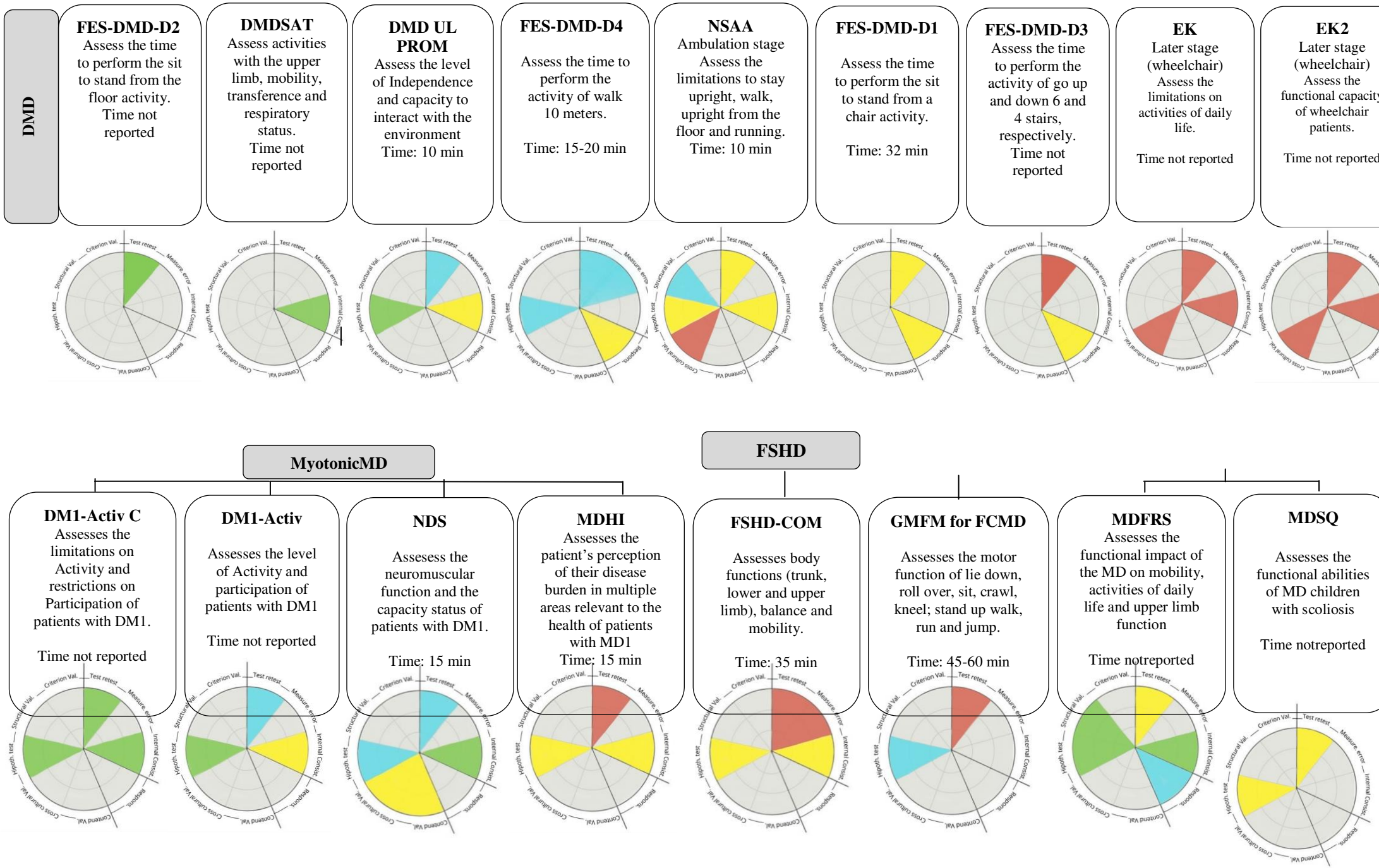

Table 1. Characteristics of the included instruments

Instrument (authors)	Construct(s)	Mode of administration	Maximum Administration Time (minutes)	(Sub)scale(s) Number of items	Response options	Score
ABILHAND questionnaire[62]	Assessment of manual skills of adult and children with NMD	Patient-reported	10	22 items	Impossible (0), difficult (1) or easy (2)	0-44
ACTIVLIM questionnaire[56, 76]	Assessment of activity limitations of patients with NMD	Patient-reported	Not reported	22 items	Impossible (0), difficult (1) or easy (2)	0-44
CAPE and CAPE –PAC[15]	Assessment of Participation and Enjoyment of children and Assessment of Preferences for Activities of Children	Patient-reported	45 - 65	55 items	1-3 1 = 'I would like to do nothing' 3 = 'I really would like to do'	0-55
CIQ[55]	Assessment of domestic, social and occupational activities	Patient-reported	5 - 10	15 items	(i) diversity; (ii) intensity; (iii) with whom; (iv) where and (v) pleasure Higher scores indicate higher community integration	0-29

DM1-Activ [16]	Assessment of the level of Activity and participation of patients with DM1	Patient-reported	Not reported	20 items	0-2 0 = enable to do 1 = able to do but it is difficult 2 = easy to do it, without difficult	1-5
DM1-ActivC [17]	For clinical use. Assessment of the limitations on Activity and restrictions on Participation of patients with DM1	Patient-reported	Not reported	25 items	0-2 0 = enable to do 1 = able to do but it is difficult 2 = easy to do it, without difficult	0-50
DMD Upper Limb PROM[18]	Assessment of the Independence level and capacity of patients to interact with the environment	Patient-reported	10	32 items 04 domain: Feeding, self-care, Family, environment leisure and communication	1-3 1 – Cannot do it 2 – Can do it with difficult 3 – Can do it easily	According to the domains
DMDSAT [58]	Assessment of the functional capacity of patients with DMD	Patient-reported	Not reported	08 items 04 domain	0-5	0-23
Egen Klassifikation (EK)[19, 63, 64]	Assessment of the functional limitation on activities of daily life of patients with DMD at later stages (loss of ambulation phase)	Patient-reported	Not reported	10 items	0-3 Higher scores indicate higher functional impairment	0-30
EK2[50, 51]	Assess the functional ability of no ambulant patients with neuromuscular disease	Patient-reported	Not reported	17 items	0-3 Higher scores indicate higher functional impairment	0-51

Upper limb functional assessment scale[18]	Assessment of the upper limb functionality	Observer-rater	15 - 23	21 items 4 dimensions	0-5	0-120
FES-DMD[21, 22, 52, 53, 68, 77, 78]	D1: Assessment of the Sit to stand from the chair activity	Observer-rater	32	1 domain	0-54	0-44
	D2: Assessment of the Sit to stand from the floor	Observer-rater	Not reported	01 domain	Stand: 0-15 Sit: 0-10	0-15 0-10
	D3: Assessment of go up and down stairs	Observer-rater	Not reported	01 domain	Lower scores indicate higher performance on the activity	0-43
	D4: Assessment of the walking activity	Observer-rater	15-20	01 domain	Lower scores indicate higher performance on the walking activity	Posture: 0-23 Balance: 0-11 Compensatory movements: 0-13
FSHD-COM[79]	Assesses body functions affected by disease	Observer-rater	35	18 items 5 sub-scales	0-4	0-72
GMFM to FCMD [25, 65]	Assessment of the motor function of patients with FCMD	Observer-rater	45 - 60	68 items 5 domain	0-8	30-50
LIFE-H [80]	Assessment of activities of daily life, social roles and the satisfaction with the performance	Patient-reported	20-30	77 items 2 sub-scales	0-9	0-10

MFM [27, 29-31, 44-46, 60, 81]	Assesses the proximal, distal and axial motor functions, based on movement, transfer and positioning commands.	Observer-rater	36	32 items 3 dimension	0-4 0 – cannot start the requested task or maintain the initial position; 1 – perform partially the task; 2 - perform partially the task or imperfect; 3 – perform the task completely; 4 – controlled movement (normal).	0-96
Short-version MFM (MFM-20)[32]	Assesses the progression of the motor function of NMD in a short manner	Observer-rater	26	20 items 3 domain	0-4 0 - cannot start the requested task or maintain the initial position; 1- perform partially the task; 2 – perform the task with compensation or slowly; 3 - perform the task completely	0-60
MDHI[33, 34, 59, 82]	Assess the perception of the patient about their disease burden in multiple areas relevant to the health of patients with myotonic dystrophy.	Patient-reported.	15	114 items 17 sub-scales	Higher scores indicate a higher disease burden	0-100
MDFRS[36]	Assessment of the functional impact of the MD on mobility, activities of daily life, upper limb function and deficiency	Patient-reported	Not reported	33 items 4 domain	0-4 1 – total dependence and 4 – Independence	0-100
MDSQ [37]	Assess the symptoms and functional abilities for DMD children with scoliosis	Patient-reported or observer-rater	Not reported	29 items	27 items: 0-5 0 – cannot do it myself 4 – it is not difficult 02 items: 0-5 0 – extremely bad 4 – no problem	Add the score and divide by the number of responses multiplied by 4

Neurological disability score (NDS) [38]	Assesses the neuromuscular function and the incapacity of patients with DM1	Patient-reported	15	21 items 4 domain	0-4 Higher scores indicate lower functional condition	0-82
Neuromuscular-Score (NM-Score)[83, 84]	Assesses the functional capacity of patients with NMD	Patient-reported or observer-rater	20-30	3 domain	0-5 0 – without impairment 4- Severe impairment.	Higher scores indicate higher functional impairment
NSAA [41, 42, 85]	Assesses the limitation of patients with DMD at the ambulation stage: stay upright, walking, stand from the chair, get up from the floor, jump and running	Observer-rater	10	17 items	0-2 2 - "Normal" – no obvious modification of the activity 1 – Activity modified but reaches the goal independently, 0 – Not able to perform the activity	0-34
PEM-CY [43]	Assesses the participation and contextual factors at home, school and community	Patient-reported	Not reported	12 items home 17 items school 16 items community	1-3 1 – Generally more difficulty; 2 - Sometimes help, sometimes it is difficulty; 3 – Sometimes help and it is not a problem	0-100

Table 2 –COSMIN classification and GRADE recommendation

Instrument	Analysis performed	Size n, (MD);%gender; Age interval; median (SD); Population	Results	COSMIN classification	Assessment of quality criteria	GRADE
ABILHAND questionnaire (62)	Content validiy	248 (78); 68% M; 6-16 age (10); DMD, BMD, LGMD, MD, FSHD and CMD	♦	Adequate	(?)	⊖(1) ⊕ ⊕⊖(1) Low[a,d]
	Hypotheses Testing (Construct validity) Comparator: HGS and ACTIVILIM	248 (78); 56% M; 16-80 (47); 6-16 age (10); DMD, BMD, LGMD, MD, FSHD and CMD	FPP: r=0.36-0.40 ACTIVILIM r=0.76	Very good	(+)	⊕ ⊕ ⊕⊖(1) Moderate[d]
ACTIVLIM questionnaire (56)	Internal Consistency	4146 (2150); 56% M; Children: 6-15 age; 11.01 (2.95), Adults: 16-92 age; 50.57 (17.74); DM1, LGMD, FSHD and DMD	♦	Doubtful	(?)	⊖(2)⊕ ⊕ ⊖(1) Very low[a,d]
ACTIVLIM questionnaire (86)	Responsiveness	132 (44); 67.43% M; 6-80 age; DMD and MD	Effect size: 0.25	Very good	(+)	⊕ ⊕ ⊕⊖(1) Moderate[d]
CAPE/PAC (15)	Internal Consistency	152 (2); 65,6; 12-18 age; DMD and other NMD	CAPE: α: 0.42-0.82 PAC: α: 0.65-0.92	Very good	(±)	⊕ ⊖(1) ⊕ ⊖(1) Low[b,d]
	Cross-cultural validity		♦	Inadequate	(?)	⊖(3) ⊕ ⊖(1) Very low[a,d]
	Reliability test-retest interval: 4 weeks		CAPE: ICC: 0.43-0.74 PAC: ICC:0.71-0.83	Adequate	(±)	⊖(1) ⊖(1) ⊕ ⊖(1) Very low[a,bd]

	Measurement Error		SEM: 4.25 SDC: 11.77	Adequate	(?)	⊖(1)⊕ ⊕ ⊖(1) Low[a,d]
CIQ[55]	Structural validity	751 (273); 18-91 age; 52; uninformed types of MD	RMSEA: 0.05	Very good	(+)	⊕ ⊕ ⊕⊖(1) Moderate[d]
	Internal Consistency		α: 0.45-0.84	Very good	(?)	⊕ ⊖(1)⊕ ⊖(1) Low [b,d]
	Hypotheses Testing (Construct validity) Comparator: General health and Mental health scale		General health: r: 0.34 Mental health scale: r:0.21	Very good	(?)	⊕ ⊖(1)⊕ ⊖(1) Low [b,d]
DM1-Activ[16]	Internal Consistency	a) Pré-fase: (163); 51,5%M; 18-69 age; 44,2 (11,7); DM1	♦	Doubtful	(?)	⊖(2)⊕⊕⊕ Low[a]
	Reliability test-retest interval: 1 year	b) Reliability: (138); 49,3 %M; 19-70 age; 46 (11,5); DM1	ICC: 0.93-0.97	Adequate	(+)	⊖(1)⊕⊕⊕ Moderate[a]
	Hypotheses Testing (Construct validity) Comparator: MRC e MIRS		MRC: ICC: 0.69 MIRS: ICC: 0.71	Very good	(+)	⊕ ⊕ ⊕⊕ High
DM1-Activ[C(17)]	Internal Consistency	a) Pré-fase: (340); 50,3%M; 18-82 age; 47.5 (12.5); DM1	♦	Very good	(?)	⊕ ⊕ ⊕⊕ High
	Reliability test-retest interval: 4 weeks	b) Reliability: (n=223); DM1	♦	Adequate	(?)	⊖(1)⊕⊕⊕ Moderate[a]
	Hypotheses Testing Comparator: DM1-Activ		DM1-Activ r: 0.91	Very good	(+)	⊕ ⊕ ⊕⊕ High
DMD Upper	Internal Consistency	(101); 100% M; 7–43 age; 15 age (7 age); DMD	♦	Doubtful	(?)	⊖(2)⊕ ⊕⊕ Low[a]
	Reliability test-retest: 7-14 days		ICC: 0.99	Adequate	(+)	⊖(1)⊕⊕⊕ Moderate[a]

Limb PROM[18]	Hypotheses Testing (Construct validity)		♦	Very good	(+)	⊕ ⊕ ⊕⊕ High
DMD SAT[58]	Internal Consistency	(186); 100% M; 5-43 age; 14 (Not described); DMD	α: 0.93	Very good	(+)	⊕ ⊕ ⊕⊕ High
EK[19]	Internal Consistency	30; 70% M; 4 e 67 age; DMD and SMA	α: 0.99	Very good	(+)	⊕⊕⊖(2)⊖(1) Very low [c,d]
	Cross-cultural validity (to Spanish population from English language)		♦	Inadequate	(?)	⊖(2) ⊕ ⊕ ⊖(1) Very low[a,d]
	Reliability test-retest: 15 days		ICC: 0.87-1.0	Adequate	(+)	⊖(1)⊕ ⊖(2) ⊖(1) Very low [a,c,d]
EK[63]	Cross-cultural validity (to Portuguese-Brazil population from English language)	(26); 100% M; 7-22 age; 12,7 (4,0); DMD	♦	Inadequate	(?)	⊖(2) ⊕ ⊕ ⊖(1) Very low[a,d]
EK[64]	Cross-cultural validity (to Portuguese-Brazil population from English language)	94 (56); 9-29 age; 100; DMD	♦	Doubtful	(?)	⊖(2) ⊕ ⊕ ⊖(1) Very low[a,d]
EK2[50]	Internal Consistency	41 (28); 6-24 age; DMD and SMA	Internal Consistency α: 0.78-0.83	Internal Consistency: Very good	Internal Consistency: (+)	⊕⊕⊖(2)⊖(1) Very low[c,d]
	Cross-cultural validity (to Turkish population from English language)		♦	Inadequate	(?)	⊕⊕⊖(2)⊖(1) Very low[c,d]
	Reliability test-retest: 1 week		ICC: 0.76-0.93	Adequate	(+)	⊕ ⊕ ⊖(1) ⊖(1) Very low[c,d]
EK2[51]	Cross-cultural validity (to Spanish population from English language)	39 (12); 59% M; 4-60 age; 16,35 (12,9); DMD and SMA	♦	Inadequate	(?)	⊕⊕⊖(2)⊖(1) Very low[c,d]

	Reliability test-retest: 4 weeks		ICC: 0.87-1.00	Very good	(+)	⊕ ⊕ ⊖(1) ⊖(1) Very low[c,d]
EESS UL[20]	Internal Consistency	10 (8); 10-16 age; 90; DMD and SMA	Internal Consistency α: 0.97	Internal Consistency: Very good	Internal Consistency: (+)	⊕ ⊕ ⊖(2) ⊖(1) Very low[c,d]
	Reliability test-retest: 2 weeks		♦	Doubtful	(?)	⊖(2) ⊕ ⊖(2) ⊖(1) Very low[a,c,d]
FES-DMD-D1[21]	Reliability test-retest: 4 weeks	30 (30); 100% M; 5-12 age; 7,4 (2,2); DMD	Reliability: ICC= 0,91 - 0,93 k= 0,88 - 0,99	Reliability: Very good	Reliability: (+)	⊕ ⊕ ⊖(2)⊖ ⊕ Low[c]
FES-DMD-D1[77]	Responsiveness 3, 6, 9 and 12 months interval	26 (26); 100% M; 5-12 age; DMD	♦	Very good	(+)	⊕ ⊕ ⊖(2)⊖ ⊕ Low[c]
FES-DMD-D2[22]	Reliability test-retest: 4 weeks	100 (100); 100% M; 5-12 age; 7,4 (2,2); DMD	ICC= 0.84 - 0.89 k= 0,80 a 1,00	Very good	(+)	⊕ ⊕ ⊕ ⊕ High
FES-DMD-D3[87]	Reliability test-retest: 4 weeks	30 (30); 5-11 age; 7,1 (2,2); DMD	ICC= 0.91 - 0.94 k= 0,79 – 1,00	Adequate	(+)	⊖(1) ⊕ ⊖(2) ⊕ Very low[a,c]
FES-DMD-D3[68]	Responsiveness 3, 6, 9 and 12 months interval	26 (26); 8.1 (1.8); DMD	◊	Very good	(+)	⊕ ⊕ ⊖(2) ⊕ Low[c]
FES-DMD-D4[53]	Responsiveness 3, 6, 9 and 12 months interval	32 (32); 9,5 age (2,7); DMD	♦	Very good	(+)	⊕ ⊕ ⊖(2) ⊕ Low[c]
	Measurement Error		SDC= 3,18 - 4,31	Very good	(?)	⊕ ⊕ ⊖(1) ⊕ Moderate[c]
FES-DMD-D4[52]	Reliability test-retest: 4 weeks	51 (51); 100% M; 5-15; 9,5 (2,7); DMD	ICC= 0,92 - 0,98	Very good	(+)	⊕ ⊕ ⊖(1) ⊕ Moderate[c]
	Measurement Error		SEM= 0,32 - 0,57	Very good	(?)	⊕ ⊕ ⊖(1) ⊕ Moderate[c]
	Hypotheses Testing (Reliability)		Age: r: 0,47 Functional timed test:	Very good	(+)	⊕ ⊕ ⊖(1) ⊕ Moderate[c]

	Comparator: Age and Functional timed test		r: 0,50			
FSHD-COM*(79)*	Internal Consistency	(41); 22-70 age; 63 FSHD	α: 0.79-0.80	Very good	(+)	⊕ ⊕ ⊖(2)⊕ Low[c]
	Reliability test-retest: 3 weeks		ICC: 0.90-0.96	Adequate	(+)	⊖(1)⊕⊖(2)⊕ Very low[a,c]
	Measurement Error		MDC: 6.67	Adequate	(?)	⊖(1) ⊕ ⊖(2)⊕ VeryLow[a,c]
	Hypotheses Testing (Convergent validity) Comparator: TUG, MMT, QMT		◊	Very good	(+)	⊕⊕⊖(2)⊕ Low[c]
GMFM for FSHD[(65)]	Hypotheses Testing (Construct validity) Comparator: Ueda's Classification and GMFM	(15); 7 ± 2 age; 53,4; FSHD	Ueda's Classification r:0.93 GMFM: r: 0.99	Very good	(+)	⊕⊕⊖(1)⊕ Moderate[c]
GMFM for FSHD[(25)]	Reliability test-retest: 6 months	(41); 0,6-24,4; 8,6 age; 46,34; FSHD	ICC: 0.97- 0.99	Adequate	(+)	⊖(1)⊕⊖(2)⊕ VeryLow[a,c]
	Hypotheses Testing Comparator: Ueda's Classification		Ueda's Classification r:0.93 HFMS r: 0.90	Very good	(+)	⊕⊕⊖(1)⊕ Moderate[c]
LIFE-H [(80)]	Reliability test-retest: 2 weeks	(28); 39.3%M; 39-75 age; 52.7 (10.01); DM1	ICC: 0.80-0.91	Very good	(+)	⊕ ⊕ ⊖(2) ⊕ Low[c]
MFM[(60)]	Criterion validity	100 (69); 18-74 age; 47 (16); FSHD; DM and other NMD	<>	Very good	(?)	⊕ ⊕ ⊕⊖(1) Moderate[d]
MFM[(28)]	Internal Consistency	303 (183); 6-62 age; 69%; DMD; BDM; LGMD; FSHD; SMD; MD; CMD and other NMD	α: 0.89-0.99	Very good	(+)	⊕⊕⊕⊖(1) Moderate [d]
	Reliability test-retest: 1 to 30 days		k: 0.51-0.80	Adequate	(?)	⊕⊕⊕⊖(1) Moderate [d]

	Hypotheses Testing (Construct validity) Comparator: VAS, Vignos and Brooke scale, FIM, HAQ CHAQ		VAS - r: 0.64-0.94 Vignos - r: 0.56-0.93 Brooke - r: 0.65- 0.87 FIM- r: 0.64-0.91 HAQ, CHAQ - r: 0.77-0.86	Very good	(+)	⊕⊕⊕⊖(1) Moderate [d]
	Responsiveness		♦	Adequate	(?)	⊕ ⊕ ⊕⊖(1) Moderate[d]
MFM[(27)]	Internal Consistency	303 (140); 6-62 age; 69%; DMD; BDM; LGMD; FSHD; MD; CMD and Other NMD	α: 0.89-0.98	Very good	(+)	⊕⊕⊕⊖(1) Moderate [d]
	Reliability test-retest: 15 to 30 days		ICC: 0.96-0.99	Very good	(+)	⊕⊕⊕⊖(1) Moderate [d]
	Hypotheses Testing (Construct validity) Comparator: VAS, Vignos and Brooke scale, FIM		VAS - r: 0.64-0.94; Vignos - r: 0.56-0.93 Brooke - r: 0.65- 0.87 FIM - r: 0.64-0.91	Very good	(+)	⊕⊕⊕⊖(1) Moderate [d]
MFM[(44)]	Structural validity	911 (698); 6-60 age; 39,18; FSHD and DM1	CFI: 0.91 - 0,96; TLI: 0.91 - 0.96; RMSEA: 0.05 - 0.7 SRMR: 1.19 - 2.10	Very good	(+)	⊕⊕⊕⊖(1) Moderate [d]
RUIZ-CORTEZ et al., 2017	Reliability test-retest: 2 weeks	48 (34); 77,1; DMD, BDM, LGMD, FSHD Other NMD (SMA and CMT)	ICC: 0.98	Adequate	(+)	⊕⊕⊕⊖(1) Moderate [d]
MFM[(30)]	Cross-cultural validity (to Portuguese-Brazil population from French language)	58 (36); 60,34% M; 6-61 age; 30,39 (not described); LGMD; FSHD; DMD; BMD; MD; MM; MCN; DM; MMC and DCTF	Cross-cultural validity: ♦	Cross-cultural validity: Inadequate	Cross-cultural validity: (?)	⊖(3)⊕⊕⊖(1) VeryLow[a,d]

	Reliability test-retest		k=0.93-1.00	Doubtful	(+)	⊕⊕⊕⊖(1) Moderate [d]
	Responsiveness	152 (100); 71,71; DMD; DMB; FSHD; LGMD; MD and CMD	♦	Very good	(?)	⊕ ⊕ ⊕⊖(1) Moderate[d]
MFM[39]	Structural validity	289 (191); CMD and MD	♦	Very good	(?)	⊕ ⊕ ⊕⊖(1) Moderate[d]
	Internal Consistency		PSI=0.59-0.92	Very good	(?)	⊕⊕⊕⊖(1) Moderate [d]
MFM[31]	Internal Consistency	105 (86); 5-14 age; 85,7; NMD: SMA, CM, CMD and HN	α: 0.95-0.99	Very good	(+)	⊕⊕⊕⊖(1) Moderate [d]
	Cross-cultural validity (to Chinese population from French language)		♦	Inadequate	(?)	⊖(3)⊕⊕⊖(1) VeryLow[a,d]
	Hypotheses Testing (Construct validity) Comparator: VAS, Vignos, Brooke		VAS - r: 0.67 - 0.94 Vignos- r: 0.25 - 0.88 Brooke - r: 0.39 - 0.64	Very good	(+)	⊕ ⊕ ⊕⊖(1) Moderate[d]
	Reliability test-retest: 1-30 days	Reliability (50)	ICC: 0.89-0.99	Adequate	(+)	⊕⊕⊕⊖(1) Moderate [d]
MFM-20[32]	Internal Consistency	88 (29); 2-7 age; 69,3; DMD, BMD, CMD and MD	α: 0.24-0.99	Very good	(±)	⊕ ⊖(1) ⊖(1) ⊖(1) Very Low[b, c, d]
	Reliability test-retest: 8-30 days	Reliability (19)	Reliability: ICC: 0.56-1.0	Reliability: Very good	Reliability: (±)	⊕ ⊖(1) ⊖(2) ⊖(1) Very Low[b, c, d]
	Hypotheses Testing (Construct validity) Comparator: VAS, Vignos and Brooke scale	Hypotheses Testing 88 (53)	VAS r: 0.56-0.86 Vignos r: 0.49-0.91 Brooke r: 0.65-0.85	Very good	(+)	⊕ ⊕ ⊖(1) ⊖(1) Low[c, d]

MDHI[59]	Cross-cultural validity (to French-Canadian from English language)	(5); 33-59 age; 60; DM1	♦	Inadequate	(?)	⊕ ⊕⊖(2)⊕ Low [c]
MDHI[33]	Internal Consistency	Internal Consistency (22)	α: 0.76-0.98	Very good	(+)	⊕⊕⊕⊕ High
	Reliability test-retest: 5-31 days	Reliability (22)	ICC: 0.69-0.97	Adequate	(?)	⊕⊖(1) ⊕⊕ Moderate [b]
	Content validity	Validade (10); 46 ± 13; 40,9; DM1	NA	Very good	(+)	⊕ ⊕⊖(2)⊕ Low [c]
	Hypotheses Testing (Construct validity) Comparator: unknown			Very good	(+)	⊕⊖(1) ⊕⊕ Moderate [b]
MDHI[34]	Internal Consistency	(60); 47,5 ± 11,2; 45; DM1	α: 0.81-0.97	Very good	(+)	⊕⊕⊕⊕ High
	Cross-cultural validity (to Italian population from English language)	Cross-cultural validity (n=11)	♦	Inadequate	(?)	⊕ ⊕⊖(2)⊕ Low [c]
	Reliability test-retest: 14 days		ICC: 0,67-0.95	Adequate	(?)	⊕⊖(1) ⊕⊕ Moderate [b]
	Hypotheses Testing (Construct validity) Comparator: Cognitive, behavioural, respiratory, visual and daytime sleepiness and quality of life tests for NMD		<>	Very good	(?)	⊕⊖(1) ⊕⊕ Moderate [b]
	Internal Consistency	(38); 45,94 ± 14,58; 55,26; DM1	α: 0.84-0.98	Very good	Consistency: (+)	⊕⊕⊕⊕ High

MDHI[82]	Cross-cultural validity (to Japanese population from English language)	(8)	♦	Very good	(?)	⊕ ⊕⊖(2)⊕ Low [c]
	Reliability test-retest: 2 weeks	(29); 47,41 ± 15,78; 65,52; DM1	ICC: 0,71-0.95	Adequate	(+)	⊕⊖(1) ⊕⊕ Moderate [b]
	Hypotheses Testing (Construct validity) Comparator: INQoL, SF-36 , WHO-QOL, FVC, MMT and HGS		<>	Very good	(?)	⊕⊖(1) ⊕⊕ Moderate [b]
MDFRS[36]	Structural validity	(n=121)	TLI: 0.943 CFI: / 0.961 SRMR:/ 0.029	Very good	(+)	⊕⊕⊕⊕ High
	Internal Consistency	(n=121)	α: 0.84-0.97	Very good	(+)	⊕⊕⊕⊕ High
	Reliability test-retest: 1 week	(51); 7-61 age; 86,27; DMD; DMB; FSHD and LGMD	ICC: 0.99	Adequate	(+)	⊖(1) ⊕⊖(1) ⊕ Low [a, c]
	Hypotheses Testing (Construct validity) Comparator: Brooke and Vignos scale, BI, muscle strength, contracture and vital capacity		<>	Very good	(+)	⊕⊕⊕⊕ High
	Responsiveness		SRM: 1.02	Very good	(+)	⊕⊕⊖(1)⊕ Moderate [c]
MDSQ[37]	Reliability test-retest: 2 weeks	(26); 5-20 age; 96,2; DMD	ICC: 0.97	Very good	(+)	⊕⊕⊖(2)⊕ Low [c]
	Hypotheses Testing (Construct validity) Comparator: ASK and PODCQ		<>	Very good	(+)	⊕⊕⊖(2)⊕ Low [c]

NDS[38]	Internal Consistency	(33); 17-71 age; DM1	α: 0.73-0.98	Very good	(+)	⊕⊕⊖(2)⊕ Low [c]
	Reliability test-retest: 1 month		ICC: 0.72-0.97	Adequate	(+)	⊖(1)⊕⊖(2)⊕ Very Low[a,c]
	Hypotheses Testing (Construct validity) Comparator: QI, VIQ e PIQ; NDS, HGS, BI, RMI, GARS, ADL, IADL, ESS		<>	Very good	(+)	⊕⊕⊖(2)⊕ Low[c]
NM- Score[88]	Cross-cultural validity (to English language) Reliability	(42); 5-19 age; 54,76; CDM	♦	Very good	(?)	⊕⊕⊖(2)⊕ Low[c]
	Hypotheses Testing (Construct validity) Comparator: Brooke scale, MFM, ACTIVLIM, Jebsen test and dinamometry		<>	Inadequate	(+)	⊖(3)⊕ ⊖(2)⊕ VeryLow[a ,c]
	Reliability test-retest interval not informed	(71)	k: 0.56-0.72	Doubtful	(?)	⊖(2) ⊖(1)⊖(1)⊖(1) VeryLow [a, b, c, d]
	Criterion validity	158 (92)	<>	Very Good	(±)	Criterion Validity: ⊕ ⊖(1) ⊕ ⊖(1) Low[b, d]
	Content validity	161	NA	Doubtful	(+)	⊖(2) ⊕ ⊕ ⊖(1) VeryLow [a, d]
NSAA[61]	Structural validity	(191); 100%M; 7-8 months; DMD	♦	Adequate	(?)	⊖(1)⊕ ⊕ ⊕ Moderate [a]
	Reliability test-retest Interval not informed		♦	Doubtful	(?)	⊕⊖(1)⊕ ⊖(1) Low[b,d]
NSAA[41]	Reliability test-retest Interval not informed	(106); 100%M; 7-12 age; DMD	ICC: 0,70 a 1	Doubtful	(+)	⊕⊖(1)⊕ ⊖(1) Low[b,d]

NSAA[42]	Cross-cultural validity (to Portuguese-Brasil population from English language)	12 (12); 100% M; 4-17 age; 12 (3,8); DMD	♦	Inadequate	(?)	⊖(3) ⊕ ⊖(2) ⊕ VeryLow[a,c]
	Internal Consistency	73 (35); 100%M; 4-15 age; 9 (2,83); DMD and healthy controls	α= 0.93	Very good	(+)	⊖(3) ⊕ ⊖(2) ⊕ Very Low[a,c]
	Reliability test-retest: 3-15 days	73 (35); 100%M; 4-15 age; 9 (2,83); DMD and healthy controls	ICC= 0,50 - 0,96	Very good	(?)	⊕⊖(1)⊕ ⊖(1) Low[b,d]
	Hypotheses Testing (Construct validity) Comparator: MFM, 6MWT and typical boys	73 (35); 100%M; 4-15 age; 9 (2,83); DMD and healthy controls	MFM: r= 0,12-0,84 TC6: r= 0,43	Very good	(+)	⊕ ⊕ ⊖(1) ⊖(1) Low[,c,d]
	Cross-cultural validity (to Turkish population from English language)	(60)	♦	Inadequate	(?)	⊖(3) ⊕ ⊖(1) ⊖(1) VeryLow[a,c,d]
	Reliability test-retest: 2 weeks	(105)	ICC: 0.67-0.93	Very good	(?)	⊕ ⊖(1) ⊕ ⊖(1) Low[b, d]
PEM-CY[43]	Internal Consistency	410 (4); 5-17; Parents of disable and healthy individuals	α= 0,67 – 0,80	Very good	(?)	⊕ ⊖(1) ⊕ ⊖(1) Low[b, d]
	Hypotheses Testing (Construct validity) Comparator: disabled children	410 (4); 5-17; Parents of disable and healthy individuals	♦	Very good	(+)	⊕ ⊕ ⊕ ⊖(1) Moderate [d]

Legend:
CAPE/PAC: Children's Assessment of Participation and Enjoyment and Preferences for Activities of Children; CIQ: Community Integration Questionnaire; DM1-Activ; DMDSAT: Functional Ability Self-Assessment Tool; DMD: Duchenne muscular dystrophy; EK: Egen Klassifikation Scale; EK2: Egen Klassifikation2 Scale; FES-DMD-D1: Functional Evaluation Scale for Duchenne muscular dystrophy Go up and sit down on a chair; FES-DMD-D2: Get up and sit on the floor; FES-DMD-D3: Go up and down stairs; FES-DMD-D4: Assessment of the walking activity; FSHD-COM: FSHD composite outcome measure; GMFM: Gross Motor function measure; GMFM for FSHD: GMFM for Fukuyama congenital muscular dystrophy; LIFE-H: Assessment of Life Habits; MFM: Motor Function Measure; MFM-20: Short-version MFM; MDHI: Myotonic Dystrophy Health Index; MDFRS: Dystrophy-specific functional rating scale; MDSQ: Muscular Dystrophy Spine Questionnaire; NDS: Neurological disability score; NMS: Neuromuscular-Score; NSAA: North Star Ambulatory Assessment; PEM-CY: Participation and Environment Measure for Children and Youth**;** MRC: Medical Research Council; MIRS: Muscular impairment rating scale; VAS: visual analogic scale, FIM: Functional Independence Measure; HAQ: Health Assessment questionnaire; CHAQ: Childhood Assessment questionnaire; QV: Quality of life; HGS: handgrip strength; 6MWT: 6-minute walk test; BI: Barthel index; M: male; DMD: Duchene's muscular

dystrophy; BMD: Becker's muscular dystrophy; LGMD: limb-girdle muscular dystrophy; MD: myotonic dystrophy; FSHD: facioscapulohumeral dystrophy; DM1: Myotonic Dystrophy type 1, SMA: spinal muscular atrophy; NMD: neuromuscular disorder, CMD: congenital muscular dystrophy; CMT: Charcot-Marie-Tooth disease; MM: mitochondrial myopathy; MMC: minicore myopathy; DCTF: congenital fiber-type distortion; ♦ important indexes not reported, ◊ multiple indexes reported, * number of articles, (+) sufficient assessment, (?) indeterminate assessment, (-) insufficient assessment; SEM: Standard Error of Measurement; SDC: Smallest Detectable Change; RMSEA: Root Mean Square Error of Approximation, α: Cronbach's alpha; ICC: intra-class correlation coefficient, k: kappa index; CFI: Comparative Fit Index; TLI: Tucker-Lewis Index, SRMR: Standardized Root Mesmo Residuals Mean; PSI: Person separation index; NA: Not applicable; INQoL: Individualized Neuromuscular Quality of Life; SF-36: 36-item Short Form Health Survey; WHO-QOL: World Health Organization Quality of Life assessment; FVC: forced vital capacity; MMT: manual muscle testing; ASK: Activities Scale for Kids; PODCQ: Pediatric Outcomes Data Collection Questionnaire; TUG: timed up and go; QMT: quantitative dynamometry; QI: intelligent quotient; VIQ: verbal inteliigence quocient; PIQ: performance intelligence quotient; NDS: neurological disability score; RMI: Rivermead mobility index; GARS: Groningen activity restriction scale, ADL: activities of daily life; ADL: activities of daily life; IADL: instrumental activities of daily life, ESS: Epworth sleepiness scale

⊕ Not downgraded factor, ⊖ Downgraded factor, (n) Downgrading level.

a Decreased to risk of bias by one level if there is serious risk of bias (several studies of questionable quality available or one study of adequate quality), two levels if there is very serious risk of bias (multiple studies of inadequate quality or one study of questionable quality available) or three levels of extreme risk of bias (only one study of inadequate quality available).

b Downgraded due to inconsistency if results are inconsistent (interpretation of results generates different interpretations).

c Decreased to inaccuracy by one level if the total sample size is less than 100 and two levels if the total sample size is less than 50.

d Decreased due to indirectness if the studies are (partially) carried out in another population or other context of use than the population or context of interest.